



**HAL**  
open science

# Utilisation de la notion de copule en tomographie

Doriano-Boris Pougaza

► **To cite this version:**

Doriano-Boris Pougaza. Utilisation de la notion de copule en tomographie. Autre [cond-mat.other].  
Université Paris Sud - Paris XI, 2011. Français. NNT : 2011PA112355 . tel-00684637

**HAL Id: tel-00684637**

**<https://theses.hal.science/tel-00684637>**

Submitted on 2 Apr 2012

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

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

## THÈSE DE DOCTORAT

Physique

“Sciences et Technologies de l'Information, des Télécommunications et des  
Systèmes”

---

Présentée par

Doriano-Boris POUGAZA

# Utilisation de la notion de copule en tomographie

---

Soutenue le 16 Décembre 2011 devant la Commission d'examen:

M.	Ali MOHAMMAD-DJAFARI	(Directeur de thèse)
M.	Jean-François BERCHER	(Co-directeur de thèse)
M.	Steeve ZOZOR	(Examineur)
M.	Emmanuel MAZER	(Rapporteur)
M.	Yannick BERTHOUMIEU	(Rapporteur)
M.	Pierre BESSIÈRE	(Examineur)
M.	Christian GENEST	(Examineur)



Thèse préparée au  
**Laboratoire des signaux et systèmes**  
*UMR8506: Univ Paris-Sud – CNRS – SUPELEC*  
3, rue Joliot curie  
91192, GIF SUR YVETTE, France

## Résumé

Cette thèse porte sur le lien entre la tomographie et la notion de copule. La tomographie à rayons X consiste à (re)construire la structure cachée d'un objet (une densité de matière, la distribution d'une quantité physique, ou une densité de loi conjointe) à partir de certaines données obtenues ou mesurées de l'objet (les projections, les radiographies, les densités marginales). Le lien entre les mesures et l'objet se modélise mathématiquement par la Transformée à Rayons X ou la Transformée de Radon. Par exemple, dans les problèmes d'imagerie en géométrie parallèle, lorsqu'on a seulement deux projections à deux angles de  $0$  et  $\pi/2$  (horizontale et verticale), le problème peut être identifié comme un autre problème très important en mathématique qui est la détermination d'une densité conjointe à partir de ses marginales. En se limitant à deux projections, les deux problèmes sont des problèmes mal posés au sens de Hadamard. Il faut alors ajouter de l'information *a priori*, ou bien des contraintes supplémentaires. L'apport principal de cette thèse est l'utilisation des critères de plusieurs entropies (Rényi, Tsallis, Burg, Shannon) permettant d'aboutir à une solution régularisée. Ce travail couvre alors différents domaines. Les aspects mathématiques de la tomographie via l'élément fondamental qui est la transformée de Radon. En probabilité sur la recherche d'une loi conjointe connaissant ses lois marginales d'où la notion de "copule" via le théorème de Sklar. Avec seulement deux projections, ce problème est extrêmement difficile. Mais en assimilant les deux projections (normalisées) aux densités marginales et l'image à reconstruire à une densité de probabilité, le lien se fait et les deux problèmes sont équivalents et peuvent se transposer dans le cadre statistique. Pour caractériser toutes les images possibles à reconstruire on a choisi alors l'outil de la théorie de probabilité, c'est-à-dire les copules. Et pour faire notre choix parmi les copules ou les images nous avons imposé le critère d'information *a priori* qui se base sur différentes entropies. L'entropie est une quantité scientifique importante car elle est utilisée dans divers domaines (en Thermodynamique, en théorie de l'information, etc). Ainsi, en utilisant par exemple l'entropie de Rényi nous avons découvert de nouvelles classes de copules. Cette thèse apporte de nouvelles contributions à l'imagerie, par l'interaction entre les domaines qui sont la tomographie et la théorie des probabilités et statistiques.

**Mots-clefs** : Tomographie, Copule, Entropie, Entropie de Tsallis, Entropie de Rényi, Entropie de Burg, Entropie de Shannon.

**Abstract**

This thesis studies the relationship between Computed Tomography (CT) and the notion of copula. In X-ray tomography the objective is to (re)construct an image representing the distribution of a physical quantity (density of matter) inside of an object from the radiographs obtained all around the object called projections. The link between these images and the object is described by the X-ray transform or the Radon transform. In 2D, when only two projections at two angles  $0$  and  $\pi/2$  (horizontal and vertical) are available, the problem can be identified as another problem in mathematics which is the determination of a joint density from its marginals, hence the notion of copula. Both problems are ill-posed in the sense of Hadamard. It requires prior information or additional criteria or constraints. The main contribution of this thesis is the use of entropy as a constraint that provides a regularized solution to this ill-posed inverse problem. Our work covers different areas. The mathematics aspects of X-ray tomography where the fundamental model to obtain projections is based mainly on the Radon transform. In general this transform does not provide all necessary projections which need to be associated with certain regularization techniques. We have two projections, which makes the problem extremely difficult, and ill-posed but noting that if a link can be done, that is, if the two projections can be equated with marginal densities and the image to reconstruct to a probability density, the problem translates into the statistical framework via Sklar's theorem. And the tool of probability theory called "copula" that characterizes all possible reconstructed images is suitable. Hence the choice of the image that will be the best and most reliable arises. Then we must find techniques or a criterion of *a priori* information, one of the criteria most often used, we have chosen is a criterion of entropy. Entropy is an important scientific quantity because it is used in various areas, originally in thermodynamics, but also in information theory. Different types of entropy exist (Rényi, Tsallis, Burg, Shannon), we have chosen some as criteria. Using the Rényi entropy we have discovered new copulas. This thesis provides new contributions to CT imaging, the interaction between areas that are tomography and probability theory and statistics.

**Keywords** : Tomography, Copula, Entropy, Tsallis'entropy, Rényi's entropy, Burg's entropy, Shannon's entropy.



# Remerciements

---

Put the text vertically centered

Je remercie d'abord toutes les personnes qui ont contribué de loin au succès de cette thèse par des conseils et leur expérience. Vous vous reconnaissez car vous êtes discrets mais vos bons conseils me parviennent toujours et surtout votre exemple et la voie à suivre que vous avez tracés pour moi.

Particulièrement tout le mérite va à mon superviseur M. Ali Mohammad-Djafari dont la rigueur scientifique, la disposition constante et surtout son côté humain est très remarquable. Je remercie aussi mon co-encadrant M. Jean-François Bercher pour les différents éclaircissements qu'il n'a cessé de m'apporter durant toutes ces années.

Un grand remerciement à tous les membres du Jury, qui se sont déplacés de très loin pour assister à ma soutenance et aussi de me prodiguer des grands conseils pour le succès de ce travail. Merci beaucoup de votre temps à lire et surtout à corriger mes multiples fautes d'orthographe et surtout de me donner beaucoup d'éclaircissements sur les différents aspects scientifiques de cette thèse.

Un remerciement sincère au Prof. Annick Weiner-Suzor et aussi un grand remerciement à tous mes Professeurs de African Institute for Mathematical Sciences (AIMS), en particulier au Prof. Neil Turok, et le Prof. Vincent Rivasseau et à tous mes amis de AIMS et de la même promotion.

Un grand merci au Prof. Didier Dacunha-Castelle dont la rencontre a été l'élément déclencheur de ma rencontre avec mon superviseur et ainsi que de ce travail de thèse.

Je ne manquerai pas de citer tous ceux qui m'ont aidé à résoudre tous mes problèmes administratifs au courant de cet exil scientifique, tout particulièrement les membres de l'administration du Laboratoire des Signaux et Systèmes (LSS), Maryvonne, Myriam, Helena, Franck, Daniel, Frederic et Celine et toute l'équipe de l'association Science Accueil (Orsay). Je n'oublierai pas de mentionner **tous les membres du LSS**, et particulièrement ceux du Groupe Problèmes Inverses (GPI) qui m'ont accueilli et m'ont aidé à me sentir à l'aise. J'ai particulièrement été chanceux d'avoir rencontré le Prof Christophe Vignat dont les conseils et les rélectures critiques de chacun de mes articles en cours m'ont ouvert des pistes. Thomas Rodet (pour son assistance informatique, et pour sa grande disponibilité légendaire), Mathieu Kowalski (merci de l'aide technique), Nicolas Gac (mon ami d'une grande gentillesse et un très bon conseiller dans tous les domaines et surtout de son soutien), Aurelia Fraysse (pour sa disposition à regarder de plus prêt tous mes calculs), Hachem Ayasso (mon ami et merci de ta gentillesse et ton aide informatique extrêmement très précieuse, finalement pour tous ses problèmes variationnels, sujet de ta recherche que tu as bien voulu appliquer au sens propre du terme dans la vie courante à mon égards, merci Hachem), Diarra (une très bonne amie et merci pour ton conseil bienveillant), Sha (mon amie qui est l'étoile chinoise). J'en profite pour saluer et encourager les prochains sur la liste : Thomas, Ning, Caifang, et enfin Leila.

Un grand merci à vous tous, mes amis : Alexandre, Sophie, Pierre-Marie, Eleni, Maja, Kevin et aussi merci à ma "dream team" depuis toujours : Eric, Davy, Aymard, Hugues, Eloi.

À toute ma famille, merci de votre soutien de “**ne jamais abandonné partout où j’irai**”, une phrase simple qui est ancrée en moi.

À ma mère, l’élue de mon coeur, merci de ton soutien et à mes trentaines de frères et soeurs qui forment la grande fratrie des Pougaza et un remerciement particulier à mes grands frères Sédar Fred, Jim Carlos, Jo-leford et tous mes petits frères et soeurs.

Merci à tous mes enseignants et condisciples de chacune des Universités que j’ai parcourues dans le monde en quête du savoir et d’un milieu propice à la recherche, tous mes enseignants de l’Université de Bangui (Rép. Centrafricaine), de Stellenbosch University, University of the Western Cape, University of the Witwatersrand (Afrique du Sud), de l’Université de Cergy Pontoise et l’Université Paris-Sud (France).

À la personification de “**Ubuntu**” qui est le “**Prisonier 46664**”, merci !

Je dédie enfin ce travail à la mémoire de mon père.

“**Singuila na ala koué, singuila na Nzapa**”.





# Table des matières

<b>Page de garde</b>	<b>1</b>
<b>Résumé</b>	<b>4</b>
<b>Tables des matières</b>	<b>13</b>
<b>1 Motivations et notations</b>	<b>15</b>
1.1 Un peu d'histoire . . . . .	15
1.2 Notations . . . . .	18
<b>2 Tomographie</b>	<b>19</b>
2.1 Tomographie en imagerie médicale . . . . .	19
2.1.1 Tomographie 2D en géométrie parallèle . . . . .	20
2.1.2 Problème direct . . . . .	22
2.1.3 Principe de l'algorithme en tomographie . . . . .	22
2.2 Tomographie avec deux projections . . . . .	25
2.3 Tomographie en géométrie parallèle . . . . .	25
2.3.1 Transformée de Radon . . . . .	26
2.3.2 Transformée à Rayons X et Transformée de Radon . . . . .	27
2.3.3 Exemples de calcul d'une transformée de Radon . . . . .	27
2.4 Conclusion . . . . .	29
<b>3 Théorie des copules</b>	<b>31</b>
3.1 Distribution multivariée . . . . .	31
3.1.1 Densité de loi conjointe de probabilité . . . . .	32
3.1.2 Fonction de répartition . . . . .	32
3.1.3 Marginales pour le cas multivarié . . . . .	32
3.1.4 Loi gaussienne multivariée . . . . .	33
3.1.5 Mélange de Gaussiennes . . . . .	34
3.1.6 Loi de Student . . . . .	34
3.1.7 Loi de Student multivariée . . . . .	36
3.2 Deux problèmes équivalents . . . . .	37

3.3	Copule et leur utilisation . . . . .	37
3.3.1	Copule bivariee . . . . .	37
3.3.2	C-volume et mesure doublement stochastique . . . . .	37
3.3.3	Copule multivariee . . . . .	40
3.3.4	Differeentes methodes de construction des copules . . . . .	40
3.4	Utilisation des copules en statistique . . . . .	43
3.4.1	Estimation au sens du maximum de vraisemblance . . . . .	43
3.4.2	Infereence des fonctions marginales (IFM) . . . . .	43
3.4.3	L'approche bayesienne . . . . .	44
3.5	Quelques points et discussion sur les copules . . . . .	45
3.6	Conclusion . . . . .	49
<b>4</b>	<b>Copules à maximum d'entropie</b>	<b>51</b>
4.1	Problèmes posés . . . . .	51
4.1.1	Problème direct . . . . .	51
4.1.2	Problème inverse . . . . .	52
4.2	Differeentes approches pour le problème inverse . . . . .	54
4.2.1	Fixer a priori une forme pour $f(x_1, x_2)$ . . . . .	54
4.2.2	Approche ad hoc de construction d'une copule . . . . .	55
4.2.3	Choisir une copule et rajouter d'autres contraintes . . . . .	56
4.2.4	Choisir une fonctionnelle $J(f)$ et l'optimiser . . . . .	56
4.3	Entropie en mecanique statistique . . . . .	57
4.3.1	Methodes proposees . . . . .	58
4.3.2	Resultats principaux . . . . .	59
4.4	Famille de copules . . . . .	62
4.5	Exemples de famille de copule bivariee . . . . .	63
4.5.1	Loi Beta . . . . .	63
4.5.2	Loi de Kumaraswamy . . . . .	64
4.5.3	Loi de puissance bilatere . . . . .	64
4.5.4	Lien avec d'autres familles de copules . . . . .	65
4.6	Les mesures de dependances . . . . .	68
4.7	Conclusions . . . . .	69

<b>5</b>	<b>Copules et tomographie discrète</b>	<b>71</b>
5.1	Introduction . . . . .	71
5.2	Copule discrète . . . . .	72
5.2.1	Produit de copules . . . . .	72
5.3	Copules décrivant les processus markoviens . . . . .	73
5.3.1	Equations de Chapman-Kolmogorov . . . . .	73
5.3.2	Équations de Chapman-Kolmogorov via les copules . . . . .	74
5.4	Matrices bistochastiques et copule discrète . . . . .	74
5.4.1	Matrices doublement stochastiques . . . . .	74
5.4.2	Copule Discrète . . . . .	75
5.5	Tomographie discrète . . . . .	76
5.5.1	Formulation du problème de reconstruction . . . . .	76
5.6	Unicité en tomographie discrète . . . . .	77
5.6.1	Méthode proposée . . . . .	77
5.6.2	Construction d’une copule par la méthode algébrique . . . . .	78
5.7	De la copule discrète à la tomographie discrète . . . . .	79
5.7.1	Algorithme pour la matrice bistochastique . . . . .	80
5.8	Représentation graphique . . . . .	81
5.9	Conclusion . . . . .	83
<b>6</b>	<b>Expérimentations numériques</b>	<b>85</b>
6.1	Lien entre la notion de copule et la tomographie . . . . .	85
6.1.1	En tomographie . . . . .	86
6.1.2	La théorie de copule . . . . .	86
6.2	Quelques représentations graphiques des nouvelles copules . . . . .	88
6.3	Screenshot of the package “Copula-tomography” . . . . .	92
6.4	Guide . . . . .	92
6.4.1	Menu . . . . .	92
6.4.2	Menu List . . . . .	93
6.5	Conclusion . . . . .	94
<b>7</b>	<b>Conclusions générales et Perspectives</b>	<b>95</b>
7.1	La classe de copules entropiques . . . . .	95

7.2	Tomographie discrète et copule discrète . . . . .	95
7.3	Extension des copules : notion de $q$ -copule . . . . .	96
7.4	Copules maximales de divergences et d'entropies . . . . .	96
7.5	Autres points . . . . .	96
7.6	Conclusions générales . . . . .	96
<b>A</b>	<b>Sur la preuve des densités de probabilité maximisant des entropies</b>	<b>99</b>
A.1	Problème d'optimisation continue . . . . .	99
A.2	Autre preuve à partir de la divergence $(a, \lambda)$ . . . . .	100
A.2.1	Remarque sur le cas indépendant . . . . .	103
A.3	Certains détails de calculs . . . . .	104
A.4	Notation . . . . .	104
A.4.1	Question . . . . .	105
A.4.2	Réponse 1 . . . . .	105
A.5	Question 2 . . . . .	106
A.6	Réponse 2 . . . . .	107
A.6.1	Règle de l'Hôpital . . . . .	107
A.7	Quelques expressions utiles . . . . .	107
A.7.1	Expression des dérivées . . . . .	107
A.7.2	Montrons que $R_q(f) = S(f)$ quand $q \rightarrow 1$ . . . . .	108
A.7.3	Calcul des limites . . . . .	108
A.7.4	Montrons que $T_q(f) = S(f)$ quand $q \rightarrow 1$ . . . . .	109
A.7.5	Question 3 . . . . .	110
	<b>Production scientifique des résultats</b>	<b>113</b>
	<b>Papier I : Copula and tomography</b>	<b>117</b>
	<b>Papier II : Lien entre la notion de copule et la tomographie</b>	<b>119</b>
	<b>Papier III : Link between copula and tomography</b>	<b>121</b>
	<b>Papier III : Maximum Entropies Copulas</b>	<b>123</b>
	<b>Papier IV : New Copulas by maximizing Tsallis or Rényi Entropies</b>	<b>125</b>





# Motivations et notations

L'un des thèmes récurrents des problèmes en mathématiques est de savoir comment retrouver une entité à partir de certaines grandeurs dont elle est dérivée. Le problème typique que nous traitons ici est la reconstruction d'une fonction de deux variables si on dispose uniquement des intégrales de cette fonction suivant les lignes droites dans l'espace formé par ces deux variables. Une classe particulière de ces problèmes, sur laquelle nous nous concentrerons, est de savoir comment déterminer une densité de probabilité conjointe de plusieurs variables connaissant ses marginales. Nous mentionnerons une des applications de ce problème dans le domaine des sciences d'imagerie. En effet on retrouve ce problème en tomographie à rayons X où on cherche une image à partir de ses projections. Dans ce chapitre nous évoquerons d'abord un peu d'historique de ces problèmes dans le domaine de la mathématique et de l'imagerie.

## 1.1 Un peu d'historique

---

En 1917, Johann Karl August Radon (1887-1956) publia un article [1] à but purement mathématique qui aura plus tard beaucoup d'applications dans divers domaines et plus particulièrement en science d'imagerie, et dans l'une de ses branches qui nous intéresse dans ce travail qui est la tomographie. En effet, la méthode analytique de reconstruction d'image en tomographie se fait via l'utilisation de l'outil mathématique développé par Radon, qui porte son nom et est appelé la transformée de Radon.

Bien que les principes mathématiques de la théorie de reconstruction d'un objet à partir de la connaissance de ses projections aient été énoncés en 1917, les applications pratiques ne seront faites qu'en 1956 [2] par Ronald Newbold Bracewell (1921-2007) dans le domaine de la radio-astronomie. Les premières applications médicales furent réalisées en 1961 par William Henry Oldendorf (1925-1992) qui avait utilisé une source de rayonnements émettant des rayons gamma et bien d'autres par la suite. La tomographie par rayons X assistée par ordinateur (appelé aussi tomodensitométrie ou plus simplement scanner) fut développée par Godfrey Newbold Hounsfield (1919-2004). Cette mise au point de la tomodensitométrie a été considérée comme la découverte la plus importante dans le domaine de la radiologie après celle des rayons X en 1895 par Wilhelm Conrad Röntgen (1845-1923) et c'est ainsi qu'elle a valu le prix Nobel de Médecine en 1979 délivré conjointement à G.N. Hounsfield et aussi à Allan MacLeod Cormack (1924-1998). La tomographie axiale assistée par ordinateur ou Computed Tomography (CT) est basée sur la détection d'un faisceau de rayons X tournant autour du patient. Cependant, à l'opposé



de la radiologie classique où le faisceau transmis est détecté et visualisé à l'aide d'un film ou d'un amplificateur de luminance, il est détecté électroniquement puis numérisé. L'image est ensuite reconstruite à l'aide d'un ordinateur et visualisée. L'acquisition de plusieurs coupes adjacentes conduit à l'information tridimensionnelle [3].

Dans le domaine des statistiques et de la théorie des probabilités, une passionnante histoire conduit à la découverte de la deuxième notion centrale qui nous intéresse dans ce rapport. Premièrement, nous devons rappeler aussi les questions posées par Maurice Fréchet (1878-1973) en 1940, dans le premier volume de l'un de ses importants travaux qu'il a écrit en trois volumes dans le domaine des probabilités [4] ayant trait à la reconstruction d'une densité jointe de probabilité qui conduiront aussi à la solution proposée par Abe Sklar en 1959 [5] à la notion de copule. Nous précisons au passage que Maurice Fréchet avait comme superviseur de thèse Jacques Salomon Hadamard (1865-1963) dont l'un des articles [6] nous intéressera dans ce travail. Les questionnements de Maurice Fréchet et bien d'autres dans le domaine de la probabilité conduiront à beaucoup de travaux de recherche que nous ne pourrons pas tous citer, jusqu'à la réponse, qui nous intéresse et qui sera donnée par Abe Sklar introduisant la notion de copule [5]. Si nous lisons aussi les travaux de Wassily Hoeffding (1914-1991) qui sont réimprimés [7, 8] il s'avère que ce dernier aurait pu faire la découverte des copules avant Abe Sklar sauf que la définition donnée était faite dans l'intervalle  $[-1/2, 1/2]$  autre que l'intervalle  $[0, 1]$  où les copules sont généralement définies, et que la publication de Wassily Hoeffding en 1940 était apparue dans un journal scientifique allemand pendant la seconde guerre mondiale à une période de trouble mondial et donc la communauté scientifique n'y avait pas fait attention, comme l'a mentionné R. Nelsen dans l'introduction de son livre [9].

Pour la première utilisation des travaux de Johann Radon [1] dans le domaine statistique et qui était d'abord faite en utilisant le lien qu'il y a avec la transformée de Fourier, cela était faite en 1936 en commençant par les travaux de H. Cramér et H. Wold [10]. Mais comme l'a mentionné Stanley R. Deans dans l'introduction de son livre [3] ce sera plutôt Alfred Rényi (1921-1970) en 1952 [11] qui va faire cette connexion dans le champs des travaux en statistique et par la suite W.M. Gilbert [12] y apportera encore plus d'éclaircissement.

Nous pourrons aussi mentionner le travail de George G. Lorentz (1910-2006) en 1949 [13] qui est connecté à celui de Alfred Rényi [11] sur les conditions qu'une fonction donnée soit reconstruite à partir de ses projections suivant certains axes (nous expliquerons en détail le terme projection). L'équivalence de ses conditions générales données dans [13] pour le cas général de l'analyse des fonctions, sera faite, mais cette fois dans un domaine restreint pour les problèmes d'analyse combinatoire sera aussi donnée par Herbert John Ryser (1923-1985) en 1957 [14] et en 1960 [15]. Il y a beaucoup d'autres travaux qui s'en suivront comme cela est indiqué dans le livre sur la tomographie discrète [16] dans le premier chapitre introductif. Bien qu'il y a eu ce lien entre les travaux de Radon et celui dans le domaine statistique il n'y a pas encore eu une connexion directe entre la transformée de Radon en tomographie et la théorie des copules en statistique. D'où notre première motivation dans ce travail.

Notre travail à la vue des rappels des travaux historiques du précédent paragraphe pourra se situer alors dans le cadre de la tomographie à faible nombre de projections, en ce sens que nous ne disposons pas de toutes les projections de la fonction à reconstruire, mais seulement de deux projections orthogonales. Et nous montrerons donc le lien qu'il

y a avec la notion de copule. Et ensuite pour le cas discret, nous montrerons aussi le lien qui existe entre la notion de copule discrète telle que définie dans le travail de W.F. Darsow, B. Nguyen et E.T. Olsen en 1992 [17] et de la tomographie discrète.

En pratique pour le cas de faible nombre de projections, le problème de la reconstruction en tomographie devient un problème inverse “mal posé”. Le terme “mal posé” est compris dans le sens donné par Jacques Salomon Hadamard [6]. Pour obtenir une solution satisfaisante à un problème inverse mal posé, une des solutions est d'utiliser de l'information a priori. Notre seconde motivation dans ce travail sera d'utiliser les informations *a priori* dont on dispose. Nous nous intéressons aux notions de copules aux entropies maximales. Pour lire une histoire passionnante sur l'origine du concept d'entropie et quelques noms des personnages ayant contribué à l'émergence de l'application de la notion d'entropie qui sera faite dans plusieurs domaines (Chimie, physique, mathématiques, théorie de l'information, etc ) dont l'un des principaux contributeurs fut notamment Ludwig Boltzmann (1844-1906) l'article [18] pourrait être consulté.

Nous nous focalisons aux expressions mathématiques de l'entropie, dont l'une la plus connue décrite en 1948 [19] par Claude Elwood Shannon (1916-2001) connu comme “le père de la théorie de l'information”. Nous utiliserons aussi l'entropie de Alfred Rényi [20, 21]. En gardant à l'esprit le thème principal de ce travail, nous recherchons toujours dès que ce sera possible le lien qu'il pourra avoir entre l'entropie utilisée et l'expression explicite de la loi de probabilité associée et enfin via le théorème de Sklar l'expression de la copule qui en découle. Il y a différentes sortes d'entropies dont on pourra retrouver les expressions dans les articles [21], [22], [23] et [24], normalement différentes familles de copules correspondantes devraient être données si nous associons les contraintes de marginales de deux projections dont nous disposons. La méthode d'apport d'information via la notion d'entropie a connu un essor sans précédent grâce aux travaux de Edwin Thompson Jaynes (1922-1998) qui publia en 1957 les deux articles [25] et [26]. La technique de maximum d'entropie décrite par E.T. Jaynes sera donc utilisée dans ce travail pour déduire l'expression explicite de la loi de probabilité jointe.

Nous espérons que ce bref rappel historique de quelques travaux permettent de mettre en lumière le cadre de recherche où ce travail se situe et à guider les différents points où ce manuscrit s'articulera autour, à savoir :

- a) Présentation des méthodes de tomographie dans le cas général et particulièrement pour le cas où on dispose de deux projections
- b) La théorie de copule et l'utilisation faite dans le domaine de statistique
- c) Les différents liens entre la tomographie et la notion de copule
- d) Quelques expérimentations numériques
- e) Des perspectives de recherche.

### 1.2 Notations

---

- $\mathbf{x} = (x_1, \dots, x_n)$
- $f$  : la densité d'une loi de probabilité jointe
- $f_i$  : la densité marginale d'une loi de probabilité
- $F_i$  : la fonction de repartition de  $f_i$
- $H$  : entropie de Shannon
- $R$  : entropie de Rényi
- $C$  : la copule
- $c$  : la densité d'une copule
- $\mathcal{H}$  : la transformée de Hilbert
- $\mathcal{R}$  : la transformée de Radon
- $D_{KL}$  : l'entropie relative ou divergence de Kullback
- $\mathcal{R}f$  : la transformée de Radon de  $f$
- $\mathcal{D}$  : la dérivée partielle de  $\mathcal{R}f$
- $\mathcal{C}$  : l'ensemble des nombres complexes  $z = a + ib$
- $\mathbb{N}$  : l'ensemble des entiers naturels
- $n!$  : la factorielle d'un entier naturel  $n$

*For a long time mathematicians felt that ill-posed problems cannot describe real phenomena and objects. However [...] the class of ill-posed problems includes many classical mathematical problems and, most significantly, that such problems have important applications [27].*

Andrei Nikolaevich Tikhonov  
(1906-1993)

# 2

## Tomographie

La tomographie est une technique d'imagerie, très utilisée dans plusieurs domaines de sciences telles que l'imagerie médicale, la géophysique, l'astrophysique et bien d'autres. C'est une technique qui permet de reconstruire le volume d'un objet à partir d'une série de mesures effectuées par tranches depuis l'extérieur (ou l'intérieur) de cet objet. Nous allons présenter ici les notions mathématiques de la tomographie dont nous aurons besoin dans les autres parties. L'un des cadres les plus connus et que nous allons considérer est le domaine de l'imagerie médicale qui utilise les rayons X. L'objet à reconstruire est en deux dimensions (2D) ou bien en trois dimensions (3D). En 3D, classiquement on reconstruisait des coupes 2D, d'où l'origine du mot tomographie. Que ce soit en 2D ou en 3D, le problème peut être étudié en deux étapes, que l'on peut décrire de la manière suivante :

- i) La modélisation directe qui décrit le lien mathématique entre les grandeurs observées et la grandeur inconnue recherchée.
- ii) L'inversion ou reconstruction permettant de retrouver ou d'estimer cette grandeur inconnue à partir de ces mesures.

En tomographie à rayons X, l'outil mathématique utilisé est la transformée de Radon.

### 2.1 Tomographie en imagerie médicale

Il y a plusieurs variantes modernes des méthodes utilisées en tomographie qui fonctionnent sur le principe consistant d'abord à collecter des données de projections suivant de multiples directions, puis ces données sont traitées par un ordinateur pour faire la reconstruction. Différents types d'acquisition du signal peuvent être utilisés dans les algorithmes de calcul similaire afin de créer une image tomographique. Il existe plusieurs types de technique de tomographie associée à un phénomène physique spécifique. Nous récapitulons dans le tableau suivant une liste des principales méthodes utilisées :

Phénomène physique	Type de tomographie
Rayons X	Tomographie à rayons X (CT)
Rayons gamma	Tomographie à émission mono-photonique (SPECT)
Annihilation électron-positron	Tomographie par émission de positons (TEP)
résonance magnétique nucléaire	Imagerie à Résonance Magnétique (IRM)
Échographie	Échographie médicale
Électrons	Microscopie électronique à transmission (MET)

Dans notre travail, nous nous concentrons sur la tomographie à rayons X (CT) qui utilise des rayons X, grâce à la transformée de Radon. Par exemple en médecine, c'est une procédure diagnostique qui utilise un équipement spécial à rayons X pour obtenir des images en coupe transversale du corps. L'ordinateur associé à cette technique affiche les images détaillées des organes, des os et autres tissus. Une application fondamentale est par exemple le diagnostic du cancer quand cette technique est utilisée pour :

- détecter ou confirmer la présence d'une tumeur ;
- fournir des informations sur la taille et la localisation de la tumeur et si elle s'est étendue ;
- une biopsie (prélèvement de cellules ou de tissus pour examen au microscope) ;
- aider à la planification d'une radiothérapie ou la chirurgie ;
- déterminer si le cancer répond au traitement.

Les exemples énumérés précédemment montrent l'importance de la tomographie.

### 2.1.1 Tomographie 2D en géométrie parallèle

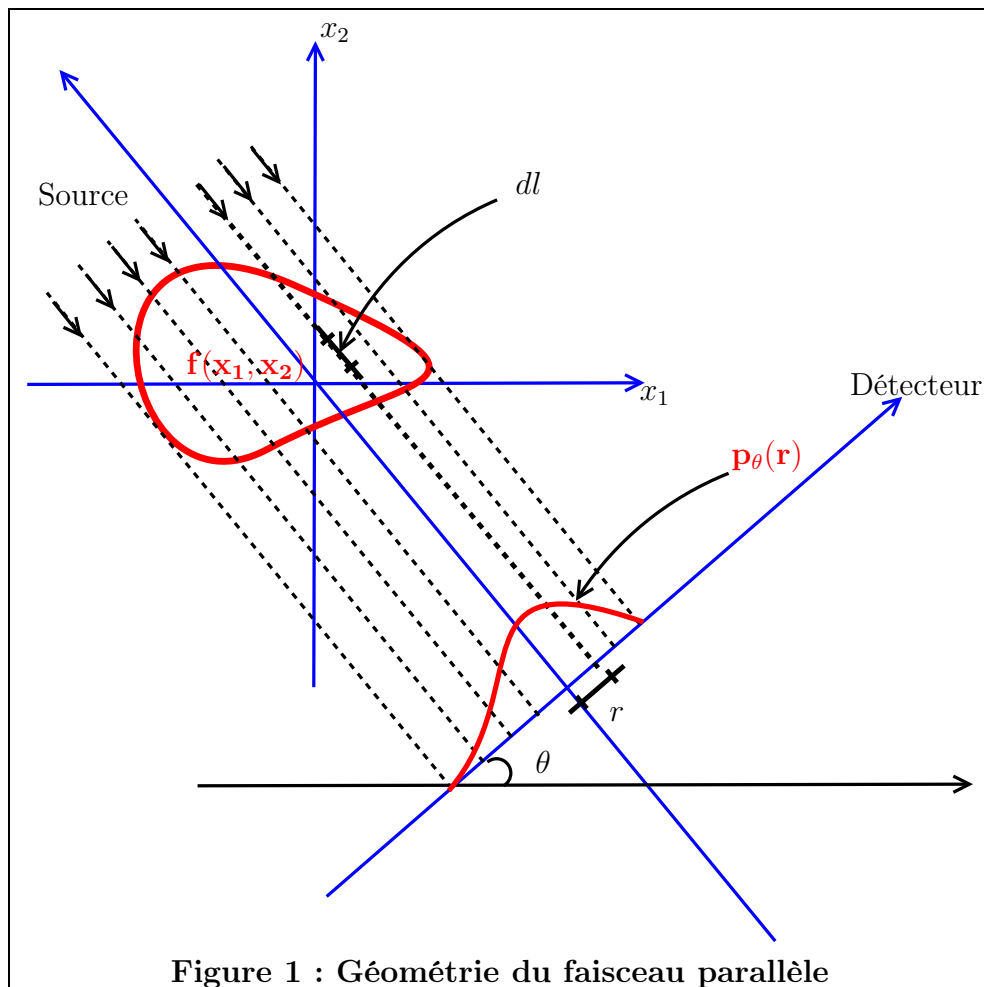
En réduisant le problème de la tomographie axiale assistée par ordinateur (traduction du mot anglais "Computed Tomography (C.T.)" dans le cas de la reconstruction d'une image à deux dimensions, donnée par une fonction  $f(x_1, x_2)$ , à partir d'un ensemble de projections à travers l'image suivant différents angles  $\theta$  dans l'intervalle  $[0, \pi)$ . Une projection  $P_\theta(r)$  est une fonction d'une variable qui est obtenue en calculant les intégrales linéiques de  $f(x_1, x_2)$  suivant des lignes parallèles passant à travers l'image, est une fonction scalaire à deux variables dans une région donnée de l'espace. Le support de l'image est souvent continu, mais un échantillonnage sur une grille de forme souvent rectangulaire est faite. Ces rectangles définissent alors un ensemble de  $n$  pixels, et l'image est représentée par un vecteur ayant comme composantes les intensités des pixels. En tomographie on ne peut pas observer l'image mais seulement des données mesurées.

L'idée principale de la CT à rayons X est d'obtenir des images de la structure interne d'un objet à l'aide de la radiographie de l'objet prise en de nombreuses directions différentes. Figure 1, pour le cas de la géométrie parallèle, les faisceaux de rayons X émis à partir d'une source suivent une trajectoire en ligne droite. Et un détecteur est placé pour faire la mesure de ses rayons X à la sortie quand ceux-ci traversent l'intérieur de l'objet  $f(x_1, x_2)$ . Chaque ligne droite est à une distance perpendiculaire  $r$  de l'origine de l'axe. L'élément de la droite noté  $dl$  est la portion sur laquelle l'intégration est faite. Pour différentes positions du détecteur en mouvement autour de l'objet suivant l'angle  $\theta$ , les données  $P(\theta, r)$  sont collectées. Pour une valeur de fixe de l'angle  $\theta$ , on obtient la valeur  $P_\theta(r)$ . L'ensemble de toutes ses valeurs mesurées constituent la donnée de projection observée du problème direct. Il faut noter que l'énergie de ses rayons X est atténuée plus ou moins en fonction de la densité de la matière rencontrée au cours de leur trajectoire (voir les travaux de Emilio Gino Segrè [28]).

L'atténuation se produit de façon exponentielle dans les tissus [29] :

$$I(r) = I_0 \exp \left( - \int_{L_{r,\theta}} f(x_1, x_2) dl \right), \quad (2.1)$$

avec  $f$  le coefficient d'atténuation au point  $(x, y)$  suivant la trajectoire des rayons  $L_{r,\theta} = \{(x_1, x_2) : r = x_1 \cos \theta + x_2 \sin \theta\}$ .



Ainsi, le modèle le plus simple est le logarithme du rapport  $I/I_0$ , avec  $I$  l'énergie mesurée sur le détecteur et  $I_0$  l'énergie émise par la source qui traverse  $f(x_1, x_2)$  le long de la droite  $L$  et pour une seule projection  $p = -\ln(I/I_0)$ . Pour un ensemble de projections, le modèle mathématique direct des données est

$$p(\theta, r) = -\ln(I/I_0) = \int_{L_{\theta,r}} f(x_1, x_2) dl. \quad (2.2)$$

L'atténuation totale est décrite pour des valeurs continues de  $0 \leq \theta < \pi$  et  $-\infty < r < \infty$ . Nous noterons  $p_\theta(r)$  quand  $\theta$  est fixe et  $r$  varie au lieu de  $p(\theta, r)$ . La projection  $p_\theta(r)$  d'un rayon à la position  $r$ , sur la projection à un angle  $\theta$ , est donnée par l'intégrale (2.2). L'équation (2.10)  $r = \langle \xi, \mathbf{x} \rangle$  et l'équation (2.5) donneront en deux dimensions, c'est-à-dire quand  $n = 2$  et pour  $\xi = (\cos \theta, \sin \theta)$ , l'ensemble des droites parallèles les unes aux autres et orientées suivant l'angle  $\theta$  et se trouvant à une distance  $r$  de l'origine du repère cartésien

$$\mathcal{R}f(\theta, r) = \int_{L_{\theta,r}} f(x_1, x_2) dl = \iint_{\mathbb{R}^2} f(x_1, x_2) \delta(r - x_1 \cos \theta - x_2 \sin \theta) dx_1 dx_2. \quad (2.3)$$

### 2.1.2 Problème direct

Le problème direct comme nous l'avons mentionné est celui consacré à l'expression mathématique des projections.

$\text{Étant données } f(x_1, x_2) \text{ trouver les projections } p_\theta(r).$

(2.4)

Il s'agit d'un problème bien posé. Selon ce que nous avons mentionné plus tôt, dans la situation où les densités de la fonction  $f$  sont connues pour toutes les positions  $(x_1, x_2)$ , les projections sont données par :

$p_\theta(r) = \iint_{\mathbb{R}^2} f(x_1, x_2) \delta(r - x_1 \cos \theta - x_2 \sin \theta) dx_1 dx_2.$

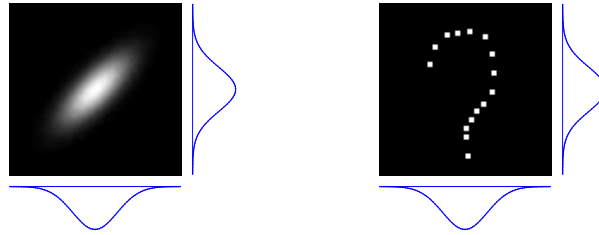
(2.5)

La relation (2.5) est connue comme la transformée de Radon (ou sinogramme) de l'objet  $f(x_1, x_2)$  dans le plan bidimensionnel, se compose de projections de rayons  $X$  ainsi que toutes les lignes possibles dans le plan. Chaque ligne a une direction spécifique et chacune des directions est identifiée par l'angle  $\theta$ .

### 2.1.3 Principe de l'algorithme en tomographie

Un problème important en statistique est la détermination d'une loi de probabilité jointe à partir de ses lois marginales. Dans le cas bidimensionnel, les lois de probabilité marginales  $f_1(x_1)$  et  $f_2(x_2)$  sont reliées à la loi jointe  $f(x_1, x_2)$  par les intégrales suivant les lignes horizontale et verticale (les deux axes  $x_1$  et  $x_2$  d'un système de coordonnées cartésiennes). Ainsi, le problème de la détermination de  $f(x_1, x_2)$  connaissant  $f_1(x_1)$  et  $f_2(x_2)$  est un problème inverse mal posé. En statistique la notion de copule est introduite

pour obtenir une solution à ce problème. Un problème similaire en tomographie à rayon  $X$  est la reconstruction d'une image  $f(x_1, x_2)$  représentant la répartition de la densité d'une quantité à l'intérieur de l'objet à partir de ses deux projections horizontale et verticale,  $f_1(x_1)$  et  $f_2(x_2)$ . Il existe aussi un grand nombre de méthodes pour de tels problèmes fondées sur la transformée de Radon. Dans ce chapitre, nous montrons les liens entre la notion de copule et celle de la tomographie à rayons  $X$  et nous voyons si on peut utiliser les méthodes d'un domaine à l'autre.



Problème directe : Donnée  $f(x_1, x_2)$  trouver  $f_1(x_1)$  et  $f_2(x_2)$   
 Problème inverse : Données  $f_1(x_1)$  et  $f_2(x_2)$  détermine  $f(x_1, x_2)$

FIGURE 2.1 – Problèmes direct et inverse

En tomographie à rayons  $X$  pour une géométrie parallèle, on a respectivement la transformée de Radon (TR) et son inverse :

$$p(\theta, r) = \iint_{\mathbb{R}^2} f(x_1, x_2) \delta(r - x_1 \cos \theta - x_2 \sin \theta) dx_1 dx_2;$$

$$f(x_1, x_2) = \frac{1}{2\pi} \int_0^\pi \int_0^\infty \frac{\frac{\partial p(\theta, r)}{\partial r}}{r - x_1 \cos \theta - x_2 \sin \theta} dr d\theta \quad (2.6)$$

qui sont les principaux outils utilisés, pour la reconstruction d'image. Nous présentons ici brièvement les principales méthodes classiques en tomographie à rayons  $X$ .

En commençant par décomposer la transformée inverse de Radon à l'aide des opérateurs suivants :

$$\text{Dérivation } \mathcal{D} : \bar{p}_\theta(r) = \frac{\partial p(\theta, r)}{\partial r},$$

$$\text{Transformée de Hilbert } \mathcal{H} : \tilde{\bar{p}}(\theta, r') = \frac{1}{\pi} \int_0^\infty \frac{\bar{p}_\theta(r)}{(r - r')} dr,$$

$$\text{Rétroprojection } \mathcal{B} : f(x_1, x_2) = \frac{1}{2\pi} \int_0^\pi \tilde{\bar{p}}(x_1 \cos \theta + x_2 \sin \theta, \theta) d\theta.$$

Ensuite, en définissant

$$P(\Omega, \theta) = \int p(\theta, r) \exp \{-j\Omega r\} dr$$



et en utilisant les propriétés de la transformée de Fourier  $\mathcal{F}$  et de la dérivation  $\mathcal{D}$  :

$$\bar{P}(\Omega, \theta) = \Omega P(\Omega, \theta)$$

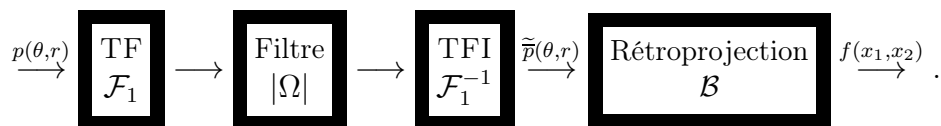
et les relations entre la transformée de Hilbert  $\mathcal{H}$  et  $\mathcal{F}$  :

$$\tilde{P}(\Omega, \theta) = \text{sign}(\Omega) \Omega \bar{P}(\Omega, \theta) = |\Omega| P(\Omega, \theta),$$

nous obtenons facilement les relations suivantes :

$$f(x_1, x_2) = \mathcal{B} \mathcal{H} \mathcal{D} p(\theta, r) = \mathcal{B} \mathcal{F}_1^{-1} |\Omega| \mathcal{F}_1 p(\theta, r) \quad (2.7)$$

qui est la méthode classique de *Rétroprojection Filtrée* (RPF) et peut se décomposer suivant le schéma



De même, si nous définissons

$$b(x_1, x_2) = \frac{1}{2\pi} \int_0^\pi p(x_1 \cos \theta + x_2 \sin \theta, \theta) d\theta \quad (2.8)$$

alors, il est montré que

$$b(x_1, x_2) = f(x_1, x_2) * h(x_1, x_2) \quad (2.9)$$

où  $*$  est l'opération de convolution en 2D et

$$h(x_1, x_2) = 1/\sqrt{x_1^2 + x_2^2}.$$

Notant que  $h(x_1, x_2)$  décroît très rapidement avec  $r = \sqrt{x_1^2 + x_2^2}$ , on peut utiliser  $b(x_1, x_2)$  comme une bonne estimation de  $f$ . C'est la méthode de rétroprojection (simple RP). En général, dans les problèmes de tomographie, si nous avons un grand nombre de projections uniformément distribuées sur un intervalle d'angle  $[0, \pi]$ , la méthode de la rétroprojection filtrée (RPF) ou même de la rétroprojection (RP) simple, fournissent de bonnes solutions. Mais, quand il y a seulement deux projections ces méthodes sont trop imprécises comme nous le montreront dans le chapitre 6.

Une description brève de l'algorithme de tomographie fonction de la manière suivante, on considère à l'entrée du code écrit dans un langage machine, un ensemble de données de projections  $\{p_{\theta_j}(r_k), j, k = 0, \dots, N - 1\}$  et on doit produire à la sortie une image approchée  $\hat{f}(x_1, x_2)$ .

- Si la résolution qui est définie par  $N$  augmente alors  $\hat{f}(x_1, x_2)$  s'approche de  $f(x_1, x_2)$ .
- S'il y a des données continues en principe l'algorithme fournit une image exacte.
- Si les données sont discrètes, on fait par exemple la supposition que ces données sont alors réparties uniformément sur le cercle polaire de rayon unitaire en grille

$$\theta_j = \frac{j\pi}{N-1}, r_k = \frac{2k - (N-1)}{N-1}$$

La supposition dans le dernier cas est que l'image  $f(x_1, x_2)$  existe dans un cercle de rayon unité et  $f(x_1, x_2)$  est nulle pour  $x_1^2 + x_2^2 > 1$ .

## 2.2 Tomographie avec deux projections

Les expressions analytiques de l'inverse de la transformée de Radon données par l'équation (2) dans les cas pair et impair donnant  $f(x_1, \dots, x_n)$  nécessitent toutes les valeurs continues de  $\xi$  et  $r$  lors de l'intégration. Mais en pratique pour les problèmes de reconstruction en tomographie on ne possède pas toutes les valeurs de  $\xi$  et  $r$ , mais le plus souvent on dispose d'un nombre faible de projections. Comme dans le industriel, en contrôle non destructif (CND) pour examiner des soudures, de matériaux composites, et faire aussi la surveillance des structures très sensibles qui pourront être des centrales nucléaires et bien d'autres exemples. Mais pour ces exemples d'applications précitées, les angles de vues ou bien des projections pour faire la reconstruction sont réduites et en générale ces problèmes se situent dans la catégorie des problèmes inverses mal posés, c'est-à-dire que le modèle mathématique associé pour résoudre le problème physique ne satisfait pas les propriétés d'un "problème bien posé" qui sont définies par Jacques Hadamard [6] :

- 1) Une solution existe
- 2) La solution est unique
- 3) Stabilité de la solution par rapport à des perturbations dans les données de problème.

Il y a plusieurs techniques pour qu'un problème mal posé puisse être résolu [30, 31].

Le problème mathématique de reconstruction en tomographie à l'aide des projections n'échappe pas à cette citation précédente et dans le cas précis de notre travail nous ne disposons que deux projections orthogonales ou de deux angles de vue alors le problème de reconstruction dans ce cas précis devient extrêmement mal posé. Mais il y a plusieurs applications de la tomographie avec deux projections [32] : par exemple quand la visualisation de nombreuses expériences tomographique est fortement limitée comme en physique des plasmas [33].

## 2.3 Tomographie en géométrie parallèle

Les illustrations suivantes montrent les modèles 3D et 2D avec les différentes projections. L'objet 3D se décompose en voxels et les projections mesurées se font suivant des plans cartésiens, contrairement au cas 2D où l'objet se décompose en pixels et les projections se font suivant des lignes.



$$p_\theta(r_1, r_2) = \int_{\mathcal{L}_{r_1, r_2, \theta}} f(x_1, x_2, x_3) dl \qquad p_\theta(r) = \int_{\mathcal{L}_{r, \theta}} f(x_1, x_2) dl$$

Problème direct (Intégrales sur des lignes) :

$$\text{En 2D : } f(x_1, x_2) \longrightarrow p_\theta(r).$$

$$\text{En 3D : } f(x_1, x_2, x_3) \longrightarrow p_\theta(r_1, r_2).$$

Problème inverse (Reconstruction d'image) :

$$\text{En 2D : } p_\theta(r) \longrightarrow f(x_1, x_2).$$

$$\text{En 3D : } p_\theta(r_1, r_2) \longrightarrow f(x_1, x_2, x_3).$$

### 2.3.1 Transformée de Radon

Le modèle mathématique utilisé est la transformée de Radon (voir l'article original publié par J. Radon en 1917 [1] et pour une traduction en Anglais de cet article voir [34]).

Si les points dans  $\mathbb{R}^n$  sont désignées par  $\mathbf{x} = (x_1, \dots, x_n)$  et les fonctions définies sur  $\mathbb{R}^n$  par  $f(\mathbf{x}) = f(x_1, \dots, x_n)$ , l'élément de volume par  $d\mathbf{x} = dx_1 dx_2 \dots dx_n$ , et l'ensemble des vecteurs unitaires de  $\mathbb{R}^n$  noté  $S^{n-1}$  ayant pour dimension  $n - 1$ ,  $\boldsymbol{\xi} = (\xi_1, \dots, \xi_n)$  est un vecteur unitaire qui définit l'orientation de l'hyperplan,  $\delta(\cdot)$  est l'impulsion de Dirac, et le produit scalaire définissant l'hyperplan par

$$\langle \boldsymbol{\xi}, \mathbf{x} \rangle = \xi_1 x_1 + \xi_2 x_2 + \dots + \xi_n x_n = r. \tag{2.10}$$

Nous choisirons  $f$  infiniment différentiable et à décroissance rapide sur  $\mathbb{R}^n$  c'est-à-dire  $\sup_{\mathbf{x}} \|\mathbf{x}^\alpha\| |\partial^\beta f(\mathbf{x})| < \infty$  pour les multi-indices  $\alpha, \beta$ , ou plus simplement  $f$  appartient à la classe de Schwartz. Il est à noter qu'il y a plusieurs définitions de la transformée de Radon  $f$  (voir [3] et les références incluses) mais toutes ces définitions sont liées et ont une forme qui est le plus souvent utilisée et donnée par :

$$\mathcal{R}f(\boldsymbol{\xi}, r) = \int f(\mathbf{x}) \delta(r - \boldsymbol{\xi} \cdot \mathbf{x}) d\mathbf{x}. \tag{2.11}$$

Considérons maintenant la solution analytique du problème inverse étant donnée la transformée de Radon. Cette solution est donnée par l'inverse de la transformée de Radon (voir [35], et aussi la preuve dans le théorème 3.53 dans [36])

1. Si  $n$  est impair, alors

$$f(\mathbf{x}) = \frac{(-1)^{\frac{n-1}{2}}}{2^n \pi^{n-1}} \int_{S^{n-1}} \mathcal{D}(\boldsymbol{\xi}, \langle \boldsymbol{\xi}, \mathbf{x} \rangle) d\boldsymbol{\xi}$$

2. Si  $n$  est pair, alors

$$f(\mathbf{x}) = \frac{(-1)^{\frac{n-1}{2}}}{2^n \pi^{n-1}} \int_{S^{n-1}} \mathcal{H}\mathcal{D}(\boldsymbol{\xi}, \langle \boldsymbol{\xi}, \mathbf{x} \rangle) d\boldsymbol{\xi},$$

avec  $\mathcal{D} = \frac{\partial^{n-1} \mathcal{R}f}{\partial r^{n-1}}$  la dérivée partielle de la transformée de Radon par rapport à sa deuxième variable.  $\mathcal{H}$  est la transformée de Hilbert de la fonction  $f$  définie sur  $\mathbb{R}$ , et est donnée par l'intégrale de la valeur principale de Cauchy de l'intégrale  $\frac{1}{\pi} \int_{-\infty}^{\infty} \frac{f(y)}{x-y} dy$  c'est-à-dire  $\mathcal{H}f(x) = \frac{1}{\pi} \lim_{\varepsilon \rightarrow 0} \int_{|x-y|>\varepsilon} \frac{f(y)}{x-y} dy$ .

### 2.3.2 Transformée à Rayons X et Transformée de Radon

Dans les dimensions supérieures, la transformée en rayons X est définie par l'intégration sur les lignes plutôt que sur des hyperplans comme dans le cas de la transformée de Radon. Cette transformée tire son nom de la tomographie à rayons X parce qu'elle représente les données de diffusion d'un balayage d'analyse tomographique à travers un milieu non homogène dont la densité est représentée par la fonction  $f$ . L'inversion de la transformée à rayons X est d'une importance pratique car elle permet de reconstruire une densité inconnue  $f$  à partir de ses données connues de diffusion.

De manière générale, si  $f$  est une fonction continue à support compact sur un espace Euclidien  $\mathbb{R}^n$ , alors la transformée à rayons X de  $f$  est la fonction  $Xf$  définie sur l'ensemble de toutes les lignes dans  $\mathbb{R}^n$  par

$$Xf(L) = \int_L f = \int f(x_0 + t\xi) dt \tag{2.12}$$

avec  $x_0$  un point initial sur la ligne et  $\xi$  est un vecteur unitaire donnant la direction de la ligne  $L$  [37].

Il faut noter que la dernière intégrale n'est pas considérée dans le sens orienté : c'est une intégrale suivant la mesure de Lebesgue à une dimension sur la ligne Euclidienne<sup>1</sup>  $L$ .

Il faut remarquer que la transformée à rayons X (2.12) est très étroitement liée à la transformée de Radon, car en dimension deux elles coïncident. D'autre part, la transformée à rayons X est aussi l'un des fondements de la géométrie intégrale moderne [38].

### 2.3.3 Exemples de calcul d'une transformée de Radon

La transformée de Radon de la fonction  $f(x_1, x_2) = e^{-x_1^2 - x_2^2}$ ,

$$\mathcal{R}f(\xi, r) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-x_1^2 - x_2^2} \delta(r - \xi_1 x_1 - \xi_2 x_2) dx_1 dx_2 \tag{2.13}$$

1. En géométrie euclidienne «La ligne droite est le plus court chemin pour aller d'un point à un autre »

En faisant le changement de variables  $u = \xi_1 x_1 + \xi_2 x_2$  et  $v = -\xi_2 x_1 + \xi_1 x_2$  tout en notant que le vecteur  $\boldsymbol{\xi} = (\xi_1, \xi_2)$  est toujours unitaire, nous avons :

$$\begin{aligned}
 \mathcal{R}f(\boldsymbol{\xi}, r) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-u^2 - v^2} \delta(r - u) \, du \, dv \\
 &= \int_{-\infty}^{\infty} e^{-v^2} \left( \int_{-\infty}^{\infty} e^{-u^2} \delta(r - u) \, du \right) \, dv \\
 &= \int_{-\infty}^{\infty} e^{-v^2} \, dv \left( e^{-r^2} \right) \\
 \mathcal{R}f(\boldsymbol{\xi}, r) &= \sqrt{\pi} e^{-r^2}.
 \end{aligned} \tag{2.14}$$

En 3D :  $f(x_1, x_2, x_3) = e^{-x_1^2 - x_2^2 - x_3^2}$ ,

les transformations

$$\begin{cases}
 u = \xi_1 x_1 + \xi_2 x_2 + \xi_3 x_3, \\
 v = -\frac{\xi_1 \xi_2}{(\xi_1^2 + \xi_2^2)^{1/2}} x_1 + (\xi_1^2 + \xi_2^2)^{1/2} x_2 - \frac{\xi_2 \xi_3}{(\xi_1^2 + \xi_2^2)^{1/2}} x_3, \\
 w = -\frac{\xi_3}{(\xi_1^2 + \xi_2^2)^{1/2}} x_1 + \frac{\xi_1}{(\xi_1^2 + \xi_2^2)^{1/2}} x_3.
 \end{cases} \tag{2.15}$$

avec  $|\boldsymbol{\xi}| = (\xi_1^2 + \xi_2^2 + \xi_3^2)^{1/2} = 1$ , conduisent à

$$\mathcal{R}f(\boldsymbol{\xi}, r) = \pi e^{-r^2}. \tag{2.16}$$

D'après [3], comme dans les cas précédents en 2D et 3D, on a la transformée de Radon  $f(x_1, \dots, x_n) = e^{-x_1^2 - \dots - x_n^2}$ , pour une valeur arbitraire  $n$  :

$$\mathcal{R}f(\boldsymbol{\xi}, r) = (\sqrt{\pi})^{n-1} e^{-r^2}. \tag{2.17}$$

La transformée de Radon est linéaire et a aussi la propriété d'homogénéité, c'est-à-dire pour tout nombre réel  $s \neq 0$  à partir de l'équation (2.11), on a :

$$\begin{aligned}
 \mathcal{R}f(s\boldsymbol{\xi}, sr) &= \int f(\mathbf{x}) \delta(sr - s\boldsymbol{\xi} \cdot \mathbf{x}) \, d\mathbf{x} \\
 &= |s^{-1}| \int f(\mathbf{x}) \delta(r - \boldsymbol{\xi} \cdot \mathbf{x}) \, d\mathbf{x}
 \end{aligned} \tag{2.18}$$

Et on note au passage pour le degré d'homogénéité  $s = -1$ , une propriété de symétrie de la transformée de Radon.

La propriété d'homogénéité (2.18), permet une importante observation : la connaissance de la transformée de Radon pour une valeur fixée  $r$  et de toutes les valeurs possibles du vecteur  $\boldsymbol{\xi}$  permet de déterminer complètement  $\mathcal{R}f$ . Par conséquent de cette observation  $\mathcal{R}f$  dépend sur  $n$  variables indépendantes si  $f$  est une fonction de  $n$  variables indépendantes. Et ces  $n$  variables indépendantes de  $\mathcal{R}f$  incluent  $r$  et les composantes du vecteur  $\boldsymbol{\xi}$  sous la contrainte que  $|\boldsymbol{\xi}| = 1$ . Comme nous l'avons déjà mentionné dans l'équation (2.11),  $\mathcal{R}f$  est définie sur  $S^{n-1} \times \mathbb{R}$ .

Les exemples suivants illustrent, comment on peut se servir de la propriété d'homogénéité et de l'observation précédentes :

Étant donné  $\zeta = s\xi$ ,  $|\zeta| = s > 0$ , il s'en suit que

$$\begin{aligned}
 \mathcal{R}f(\zeta, r) &= \mathcal{R}f(s\xi, r) \\
 &= |s^{-1}| \int f(\mathbf{x}) \delta\left(\frac{r}{s} - \xi \cdot \mathbf{x}\right) d\mathbf{x} \\
 &= |s^{-1}| \mathcal{R}f\left(\xi, \frac{r}{s}\right) \\
 \mathcal{R}f(\zeta, r) &= |\zeta^{-1}| \mathcal{R}f\left(\frac{\zeta}{|\zeta|}, \frac{r}{|\zeta|}\right).
 \end{aligned} \tag{2.19}$$

Et par exemple, à partir de l'équation (2.17), nous obtiendrons :

$$\mathcal{R}f(\zeta, r) = \frac{(\sqrt{\pi})^{n-1}}{|\zeta|} \exp\left(-\frac{r^2}{|\zeta|^2}\right), \tag{2.20}$$

avec  $|\zeta| = (\zeta_1^2 + \dots + \zeta_n^2)^{1/2}$ .

## 2.4 Conclusion

Dans ce chapitre nous avons montré succinctement l'importance capitale de la tomographie dans un domaine vital qui est l'imagerie médicale. Nous avons énuméré les différentes modalités (IRM, tomographie axiale, Échographie, etc.). Pour le cas de la tomographie à Rayons X en géométrie parallèle, nous avons montré le modèle mathématique utilisée et aussi le problème inverse de reconstruction d'un objet. À l'aide de quelques exemples de fonctions, nous avons montré les propriétés et aussi comment calculer la transformée de Radon qui est l'outil principal en tomographie.



# 3

## Théorie des copules

Ce chapitre a pour objectif de fournir une introduction à la notion de copule. Nous commencerons par les définitions des distributions multivariées et leurs marginales avec quelques exemples typiques. Ensuite naturellement nous exposerons les définitions liées aux copules, d'abord le cas bivarié et ensuite le cas multivarié. Ensuite, nous exposerons d'une manière synthétique les différentes méthodes de construction des copules et c'est là que nous pourrions introduire une de notre contribution essentielle qui est l'utilisation des critères entropiques. Ceci va être ensuite exposé en détail au chapitre 4.

Pour être complet, nous exposerons aussi l'utilisation des copules en statistique en deux points complémentaires :

- La génération des variables dépendantes en utilisant des copules.
- L'estimation des paramètres d'une copule à partir des observations

Nous finirons ce chapitre par quelques remarques et discussions ouvertes à propos de la construction et l'utilisation des copules d'un point de vue mathématique. Ceci nous conduira tout naturellement à la partie essentielle et des résultats originaux de cette thèse qui seront exposés dans les chapitres qui suivront.

Nous donnons quelques définitions et propriétés générales qui sont classiques en statistique. On admettra bien sûr que les conditions pour la dérivation, l'intégration sont garanties et ensuite nous introduisons la notion des copules et leur utilisation dans le domaine de traitement de données statistiques.

### 3.1 Distribution multivariée

---

Nous donnons d'abord un bref résumé de la définition et des propriétés d'une distribution multivariée. Car à chaque distribution multivariée, on peut associer une copule par le théorème de Sklar qui sera aussi donné ici. Nous supposons, en effet quand c'est possible que les variables sont continues.



### 3.1.1 Densité de loi conjointe de probabilité

Soient  $X_1, \dots, X_n$ ,  $n$  variables continues<sup>1</sup>. La **densité de loi de probabilité conjointe** de  $X_1, \dots, X_n$ , notée  $f_{(X_1, \dots, X_n)}(x_1, \dots, x_n)$ , est la fonction  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  tel que pour n'importe quel domaine  $D \subset \mathbb{R}^n$  dans l'espace de  $n$  dimension contenant les valeurs des variables  $X_1, \dots, X_n$ , la probabilité qu'une réalisation de l'ensemble de ces variables tombe dans le domaine  $D$  est

$$\mathbb{P}(X_1, X_2, \dots, X_n \in D) = \int_D f(x_1, \dots, x_n) dx_1 \dots dx_n. \quad (3.1)$$

**Proposition 3.1.1.** *Les deux propriétés suivantes sont satisfaites :*

- $f(x_1, \dots, x_n) \geq 0$ ,  $\forall (x_1, \dots, x_n)$ ,
- $\int_{\mathbb{R}^n} f(x_1, \dots, x_n) dx_1 \dots dx_n = 1$ .

### 3.1.2 Fonction de répartition

Une fonction de répartition (f.d.r.) dans le cas de variables prenant valeur continue est définie par :

$$F(x_1, \dots, x_n) = \int_{-\infty}^{x_1} \dots \int_{-\infty}^{x_n} f(y_1, \dots, y_n) dy_1 \dots dy_n.$$

Il y a une relation entre la densité de probabilité (si elle existe) et la fonction de répartition :

$$f(x_1, \dots, x_n) = \frac{\partial^n F(x_1, \dots, x_n)}{\partial x_1 \dots \partial x_n}. \quad (3.2)$$

### 3.1.3 Marginales pour le cas multivarié

**Définition 3.1.3.1.** *Soit  $\{X_i : i \in I\}$  un ensemble de variables aléatoires telle que le sous-ensemble  $I \subset \{1, 2, \dots, n\}$ . Si  $f(x_1, \dots, x_n)$  est la densité conjointe (sous les hypothèses adéquates) définie sur  $\overline{\mathbb{R}}^n$ . La fonction de répartition marginale  $F_{\{X_i : i \in I\}}$ , est obtenue en sommant (pour les variables discrètes)*

$$F_{\{X_i : i \in I\}}(x_i) \equiv F_i(x_i) = \sum_{\{x_i : i \notin I\}} f_{X_1, \dots, X_n}(x_1, \dots, x_n), \quad (3.3)$$

ou en intégrant (pour les variables continues)

$$F_{\{X_i : i \in I\}}(x_i) \equiv F_i(x_i) = \int_{\{x_i : i \notin I\}} f_{X_1, \dots, X_n}(y_1, \dots, y_n) \prod_{\{y_i : i \notin I\}} dy_i \quad (3.4)$$

sur toutes les valeurs des autres variables.

---

1. Brièvement une variable aléatoire multivariée est une fonction définie sur l'ensemble des résultats possibles d'une expérience aléatoire ou bien éventualités et dans le cas considéré ici les variables aléatoires sont réelles.

De même on a les densités marginales

$$f_i(x_i) = \int_{(x_1, \dots, x_n) / \{x_i: i \notin \{1, \dots, n\}\}} f(x_1, \dots, x_n) dx_1 \dots dx_{i-1} dx_{i+1} \dots dx_n. \quad (3.5)$$

On notera par  $F(x_1, x_2)$  la fonction de répartition bivariée, par  $f(x_1, x_2)$  sa densité de probabilité par  $F_1(x_1)$  et  $F_2(x_2)$  ses fonctions de répartition marginales (f.d.r.m.) et  $f_1(x_1), f_2(x_2)$  leurs densités de probabilités respectives et les relations classiques suivantes qui lient ces différentes fonctions<sup>2</sup> :

$$F(x_1, x_2) = \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} f(s, t) ds dt; f(x_1, x_2) = \frac{\partial^2 F(x_1, x_2)}{\partial x_1 \partial x_2};$$

$$F_1(x_1) = \int_{-\infty}^{x_1} f_1(s) ds = F(x_1, \infty); F_2(x_2) = \int_{-\infty}^{x_2} f_2(t) dt = F(\infty, x_2);$$

$$f_1(x_1) = \frac{\partial F_1(x_1)}{\partial x_1} = \int f(x_1, x_2) dx_2; f_2(x_2) = \frac{\partial F_2(x_2)}{\partial x_2} = \int f(x_1, x_2) dx_1.$$

**Remarque 1.** Les définitions précédentes montrent le cas classique où les marginales sont données pour une seule variable  $x_i$ , mais on peut aussi étendre les définitions dans les équations (3.3), (3.3) et (3.5) aux lois marginales multivariées, en interprétant  $x_i$  comme un vecteur de plusieurs variables, c'est-à-dire en redéfinissant  $x_i \equiv \mathbf{x}_i = (\mathbf{x}_1, \dots, \mathbf{x}_n)$ .

### 3.1.4 Loi gaussienne multivariée

Le modèle de mélanges gaussiens (MG) est l'un des modèles statistiques les plus utilisés. C'est un modèle qui fournit un cadre de travail convenable pour l'estimation des densités, car n'importe quelle densité continue peut être approchée avec une précision arbitraire par un mélange fini de densités gaussiennes. Par exemple avec 10000 observations, une densité log normale peut être approchée par un mélange de 30 lois normales, en comparaison avec un estimateur par noyaux<sup>3</sup> qui utilise 10000 lois normales (voir [40] pour une preuve). Vu l'importance des lois de mélange dites normales ou gaussiennes, nous rappelons aussi sa forme qui est définie à travers une loi de densité gaussienne.

Soit  $\mathbf{X} = (X_1, \dots, X_n)^T$  un vecteur de variable aléatoire de dimension  $n$  qui est distribuée suivant une loi gaussienne, alors la densité de la fonction de probabilité est

$$f(\mathbf{x} \mid \boldsymbol{\mu}, \boldsymbol{\Sigma}) = \frac{1}{(2\pi)^{n/2} |\boldsymbol{\Sigma}|^{1/2}} \exp\left(-\frac{1}{2} (\mathbf{x} - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu})\right), \quad (3.6)$$

où

2. Voir par exemple [39] à la page 11, en tenant compte de la liste des erreurs d'impression (errata) fournie par l'auteur.

3.  $\hat{f}_h(x) = \frac{1}{nh} \sum_{i=1}^n K\left(\frac{x-x_i}{h}\right)$ , "h" un paramètre nommé "fenêtre", qui régit le degré de lissage de l'estimation et "K" est choisi comme étant la densité d'une fonction gaussienne standard.

- $\mathbf{x} = [x_1, \dots, x_n]^T$
- $\boldsymbol{\mu} = [\mu_1, \dots, \mu_n]^T$ , est le vecteur constitués des valeurs moyennes  $\mu_i = E(X_i)$ ,
- $\boldsymbol{\Sigma}$  et la matrice de covariance, de dimension  $n \times n$  non singulière, définie positive sur l'ensemble des réels avec des éléments  $\Sigma_{ij} = E[(X_i - \mu_i)(X_j - \mu_j)]$ .

Quand  $\mathbf{X}$  est répartie suivant une loi normale multivariée, nous utilisons la notation suivante :

$$\mathbf{X} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \text{ ou } f(\mathbf{x} \mid \boldsymbol{\mu}, \boldsymbol{\Sigma}) = \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma}).$$

On peut alors déduire simplement l'expression de la fonction de repartition, qui est donnée par

$$\Phi(\mathbf{X}) = \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_n} \frac{1}{(2\pi)^{n/2} |\boldsymbol{\Sigma}|^{1/2}} \exp\left[-\frac{1}{2} (\mathbf{x}_2 - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1} (\mathbf{x}_2 - \boldsymbol{\mu})\right] dy_1 \dots dy_n. \quad (3.7)$$

## Cas 2D

Dans le cas bivarié, la loi de la densité de probabilité est :

$$f(x_1, x_2) = \frac{1}{2\pi\sigma_{x_1}\sigma_{x_2}\sqrt{1-\rho^2}} \exp\left(\frac{-1}{2(1-\rho^2)} \left(\frac{(x_1 - \mu_1)^2}{\sigma_{x_1}^2} + \frac{(x_2 - \mu_2)^2}{\sigma_{x_2}^2} - \frac{2\rho(x_1 - \mu_1)(x_2 - \mu_2)}{\sigma_{x_1}\sigma_{x_2}}\right)\right),$$

où  $\rho$  est la corrélation entre  $x_1$  et  $x_2$ , et  $\boldsymbol{\Sigma} = \begin{bmatrix} \sigma_{x_1}^2 & \rho\sigma_{x_1}\sigma_{x_2} \\ \rho\sigma_{x_1}\sigma_{x_2} & \sigma_{x_2}^2 \end{bmatrix}$ . Dans ce cas, il est évident que  $f(x_1) = \mathcal{N}(\mu_{x_1}, \sigma_{x_1})$  et  $f(x_2) = \mathcal{N}(\mu_{x_2}, \sigma_{x_2})$ .

### 3.1.5 Mélange de Gaussiennes

Pour un mélange de lois gaussiennes, chacune des  $K$  composantes  $X_k$ , est telle que  $X_k \sim \mathcal{N}(\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)$ , avec une composante de masse de probabilité  $\alpha_k$ . Cependant, les densités de mélange de loi de probabilité ont la forme suivante dans le cas des variables discrètes (il suffit de faire un changement du symbole de sommation en symbole d'intégrale pour le cas des variables continues) :

$$f(\mathbf{x}) = \sum_{k=1}^K \alpha_k \mathcal{N}(\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k), \quad (3.8)$$

où  $0 \leq \alpha_k \leq 1$ ,  $\sum_{k=1}^K \alpha_k = 1$ , et  $\mathbf{x} = (x_1, \dots, x_n)$ .

### 3.1.6 Loi de Student

La loi de Student (ou  $t$ -distribution) avec  $\nu$  degrés de liberté a la loi de densité de probabilité :

$$f(t) = \frac{1}{\sqrt{\nu} B\left(\frac{\nu}{2}, \frac{1}{2}\right)} \left(1 + \frac{t^2}{\nu}\right)^{-\frac{\nu+1}{2}}, \quad (3.9)$$

où  $B(\alpha, \beta)$  est la *fonction bêta* qui est définie comme.

$$B(p, q) = \int_0^1 x^{p-1}(1-x)^{q-1} dx$$

pour des nombres réels  $p, q > 0$ . Nous noterons par  $X \sim t(\nu)$ . Nous pouvons aussi écrire la *fonction bêta* en utilisant *fonction gamma*,

$$\Gamma(z) = \int_0^\infty e^{-t} t^{z-1} dt$$

pour  $z \in \mathcal{C}$  avec  $\Re(z) > 0$ , et continuité analytique pour le reste du plan complexe, à l'exception pour les entiers naturels négatifs, où il y a des pôles simples. Une autre définition semblable est

$$\Gamma(z) = \frac{e^{-\gamma z}}{z} \prod_{n=1}^{\infty} \left(1 + \frac{z}{n}\right)^{-1} e^{z/n},$$

ou  $\gamma$  est la constante d'Euler.

Les fonctions *bêta* et *gamma* ([41]) sont telles que :

- $B(p, q) = \frac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)}$  pour  $p, q \in \mathcal{C}$  avec le côté droit de l'égalité bien défini.
- $B(p, q) = B(q, p)$  et  $B\left(\frac{1}{2}, \frac{1}{2}\right) = \pi$ .
- $\Gamma(p) = (p-1)\Gamma(p-1)$  et  $\forall n \in \mathbb{N}, \Gamma(n) = (n-1)!$

Par conséquent la Loi de Student est donnée par :

$$f(x) = \frac{\Gamma(\frac{\nu+1}{2})}{\sqrt{\nu\pi} \Gamma(\frac{\nu}{2})} \left(1 + \frac{x^2}{\nu}\right)^{-\frac{\nu+1}{2}}. \quad (3.10)$$

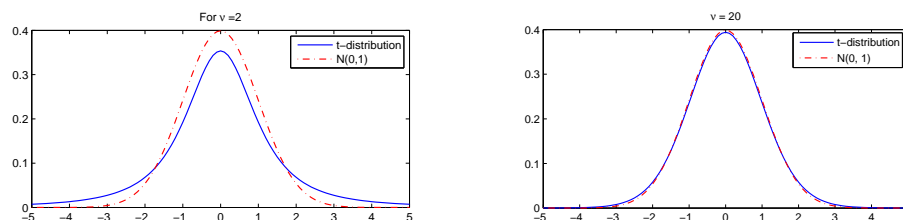
**Proposition 3.1.2.** Avec les propriétés suivantes

- $t(1) = \text{Cauchy}(0, 1)$ , est la loi de Cauchy avec les paramètres 0 et 1.
- $t(\nu) \sim \mathcal{N}(0, 1)$  quand  $\nu \rightarrow \infty$ .
- Si  $X \sim t(\nu)$ ,  $E[|X|^k]$  existe si et seulement si  $k < \nu$ .

La dernière propriété signifie qu'il n'y a pas de moyenne pour la loi de Student quand  $\nu = 1$ . Et pour  $k = 1$ , avec  $\nu > 1$ ,  $E[X] = 0$ ; et pour le cas  $\nu > 2$ , la variance de  $X$  est égale à  $\frac{\nu}{\nu-2}$ . Si  $X_1, \dots, X_n$  sont des variables aléatoires d'une loi gaussienne avec la moyenne  $\mu$  et la variance  $\sigma^2$ . Si nous notons  $\bar{\mu}, \bar{\sigma}^2$  respectivement les échantillons de la moyenne  $\bar{\mu} = \frac{1}{n} \sum_{k=1}^n x_i$  et de la variance  $\bar{\sigma}^2 = \frac{1}{n} \sum_{k=1}^n (x_i - \bar{\mu})^2$ , alors

$$\frac{\bar{\mu} - \mu}{\bar{\sigma}} \sqrt{\nu} \sim t(\nu - 1). \quad (3.11)$$

Les courbes suivantes illustrent les cas  $\nu = 10, \nu = 20$ , qui nous permettent de constater que si  $\nu$  augmente, asymptotiquement la loi de student atteint la loi normale gaussienne.



### 3.1.7 Loi de Student multivariée

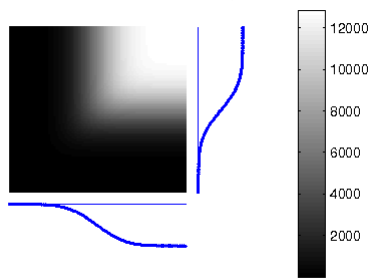
Soit  $\mathbf{X}$  un vecteur de  $n$  variables suivant une loi de Student avec  $\nu$  degrés de liberté et avec la moyenne  $\boldsymbol{\mu}$ . D'après les propriétés précédentes dans le cas univarié, nous posons la moyenne  $\mu > 1$  et le degré de liberté  $\nu > 2$ , la matrice de covariance est  $\frac{\nu}{\nu-2}\boldsymbol{\Sigma}$ .

Alors  $\mathbf{X}$  a la forme

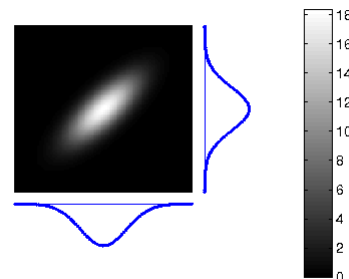
$$\mathbf{X} = \boldsymbol{\mu} + \frac{\sqrt{\nu}}{\sqrt{S}}\mathbf{Z}$$

où  $\mathbf{Z} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma})$  et  $S$  est indépendamment distribuée de  $\mathbf{Z}$  et avec une distribution  $\chi^2_\nu$ , alors la loi de Student multivariée est donnée par

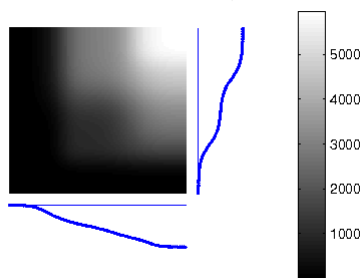
$$f(\mathbf{x}) = \frac{\Gamma\left(\frac{\nu+n}{2}\right)}{(\nu\pi)^{n/2}} \Gamma\left(\frac{\nu}{2}\right) |\boldsymbol{\Sigma}|^{-1/2} \left(1 + \frac{1}{\nu}(\mathbf{x} - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1}(\mathbf{x} - \boldsymbol{\mu})\right)^{-\frac{\nu+n}{2}}. \quad (3.12)$$



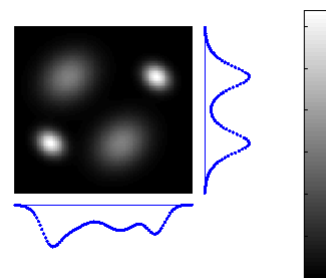
f.d.r. et les marginales



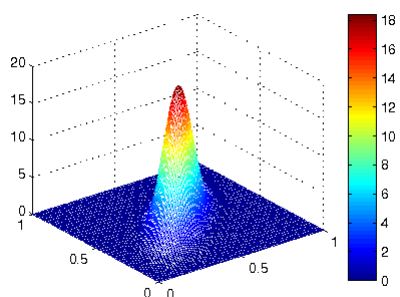
Densité gaussienne et ses marginales



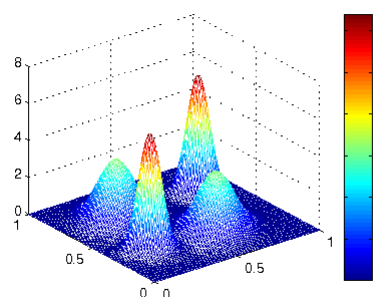
MG de f.d.r. et les marginales



Densité de MG et les marginales



Densité gaussienne



MG

## 3.2 Deux problèmes équivalents

D'une manière générale les relations entre la densité de la loi conjointe inconnue (ou l'objet)  $f(x_1, \dots, x_n)$  à reconstruire connaissant ses densités marginales ( ou les projections , c'est à dire les quantités physiques mesurées)  $f_i(x_i)$  se posent de la manière suivante :

- Problème direct : Étant donnée  $f(x_1, \dots, x_n)$  trouver  $f_i(x_i)$
- Problème inverse : Étant données les  $f_i(x_i)$  trouver  $f(x_1, \dots, x_n)$ .

Pour le problème direct si les  $f_i(x_i)$  sont des densités marginales (assimilées aux projections en nombre suffisant et continues ) alors le problème direct est la solution de l'équation (3.5).

Mais comment résoudre le problème inverse qui est celui de trouver la loi conjointe (ou bien l'objet ) quand les densités marginales (les projections ou bien des données rayons X)  $f_i(x_i)$  sont données?

Dans les sections suivantes, nous allons voir comment appréhender ce problème en présentant l'outil adéquat pour l'estimation d'une densité conjointe à partir de ses marginales.

## 3.3 Copule et leur utilisation

### 3.3.1 Copule bivariée

Une copule bivariée  $C$  est une fonction de  $[0, 1]^2$  dans  $[0, 1]$  telle que :

- $\forall (u_1, u_2) \in [0, 1]^2, C(u_1, 0) = C(0, u_2) = 0,$
- $\forall (u_1, u_2) \in [0, 1]^2, C(u_1, 1) = u_1, C(1, u_2) = u_2,$
- $\sum_{k=1}^2 \sum_{l=1}^2 (-1)^{(k+l)} C(u_1^{(k)}, u_2^{(l)}) = C(u_1^{(1)}, u_2^{(1)}) - C(u_1^{(1)}, u_2^{(2)}) - C(u_1^{(2)}, u_2^{(1)}) + C(u_1^{(2)}, u_2^{(2)}) \geq 0$  avec  $0 \leq u_i^{(1)} \leq u_i^{(2)} \leq 1$  pour  $i = 1, 2.$

Notons tout de suite que  $C(u_1, u_2)$  est une fonction de repartition d'une loi de densité de probabilité bivariée  $c$  et que dans le cas continu, on a  $C(u_1, u_2) = \int_0^{u_1} \int_0^{u_2} c(s, t) ds dt.$

Les deux premières propriétés i) et ii) sont les conditions aux limites de  $C$ . Et la dernière propriété iii) est l'extension bidimensionnelle de la notion classique utilisée pour la monotonie d'une fonction ( ici c'est la croissante d'une fonction ). Par conséquent une copule  $C(u_1, u_2)$  est une fonction 2-croissante. La définition dans le cas de fonction multivariée est la suivante.

### 3.3.2 C-volume et mesure doublement stochastique

Soient  $\mathbf{a} = (a_1, a_2, \dots, a_n)$  et  $\mathbf{b} = (b_1, b_2, \dots, b_n)$ . Soit  $B$  un pavé de dimension  $n$  i.e  $B$  est égale au produit cartésien  $[a_1, b_1] \times [a_2, b_2] \times \dots [a_n, b_n]$ . Et soit  $\mathbf{c} = (c_1, c_2, \dots, c_n)$

les sommets de  $B$ , c'est-à-dire que  $c_k$  est égale à  $a_k$  ou  $b_k$ , et que  $\mathbf{a} \leq \mathbf{b}$  équivaut à  $a_k \leq b_k \quad \forall k = 1 \dots n$ .

**Définition 3.3.2.1.** Soient  $S_1, S_2, \dots, S_n$  les sous-ensembles non vides de la droite numérique achevée  $\overline{\mathbb{R}} = [-\infty, +\infty]$ , et  $C$  une fonction réelle de  $n$ -uplet, c'est-à-dire tel que son domaine  $\text{Dom}(C)$  est un sous-ensemble de  $\overline{\mathbb{R}}^n$  et dont l'ensemble de valeurs prises de  $C$  noté  $\text{Ran}(C)$ , est un sous-ensemble de  $\mathbb{R}$ , et que  $\text{Dom}(C) = S_1 \times S_2 \times \dots \times S_n$ . Pour le pavé  $B = [\mathbf{a}, \mathbf{b}]$  tels que tous ses sommets sont dans  $\text{Dom}(C)$  le  $C$ -volume de  $B$  est donné par

$$V_C(B) = \sum_{\mathbf{c} \in B} \text{sign}(\mathbf{c})C(\mathbf{c}), \quad (3.13)$$

où  $\text{sign}(\mathbf{c}) = 1$ , si  $c_k = a_k$  pour un nombre  $k$  pair, et  $\text{sign}(\mathbf{c}) = -1$ , si  $c_k = a_k$  pour  $k$  impair.

On peut aussi exprimer le  $C$ -volume du pavé  $B$  en terme de la  $n$ -ième différence de  $C$  sur  $B$

$$V_C(B) = \Delta_{\mathbf{a}}^{\mathbf{b}} C(\mathbf{t}) = \Delta_{a_n}^{b_n} \Delta_{a_{n-1}}^{b_{n-1}} \dots \Delta_{a_2}^{b_2} \Delta_{a_1}^{b_1} C(\mathbf{t}),$$

avec la  $n$ -ième différence de premier ordre de  $C$  étant

$$\Delta_{\mathbf{a}_k}^{\mathbf{b}_k} C(\mathbf{t}) = C(t_1, \dots, t_{k-1}, b_k, t_{k+1}, \dots, t_n) - C(t_1, \dots, t_{k-1}, a_k, t_{k+1}, \dots, t_n).$$

Pour  $n = 2$  une copule  $C$  induit une  $C$ -mesure  $V_C$  sur  $[0, 1]^2$ , et c'est une mesure appelée *mesure doublement stochastique* par la correspondance

$$V_C(B) = \Delta_{a_2}^{b_2} \Delta_{a_1}^{b_1} C(\mathbf{t}) = [C(b_1, b_2) - C(b_1, a_2)] - [C(a_1, b_2) - C(a_1, a_2)],$$

qui est une mesure de rectangle (ou bien simplement un pavé de dimension 2). Réciproquement, pour n'importe quelle mesure doublement stochastique  $V$ , il existe une copule unique  $C_V$  définie par  $C_V(u_1, u_2) = V([0, u_1] \times [0, u_2])$  ([9] page 26)

Toute la théorie sur les copules a son fondement basé sur le théorème suivant :

**Théorème 3.3.1** (Théorème de Sklar (voir [42] pour la preuve)). Soit  $H$  une fonction de répartition avec ses fonctions de répartition de ses lois marginales  $F$  et  $G$ . Alors il **existe** une copule  $C$  telle que :

$$H(x_1, x_2) = C(F(x_1), G(x_2)). \quad (3.14)$$

Si de plus les fonctions de lois marginales sont continues, alors la copule  $C$  est **unique**, et est donnée par

$$C(u_1, u_2) = H(F^{-1}(u_1), G^{-1}(u_2)). \quad (3.15)$$

Autrement la copule  $C$  est déterminée de manière unique sur  $\text{Ran } F \times \text{Ran } G$ . **Réciproquement**, étant données des fonctions de répartition  $F$  et  $G$ , avec  $C$  une copule quelconque, alors la fonction  $H$  définie par (3.14) est une fonction de répartition bivariée avec les marginales  $F$  et  $G$ .

**Théorème de Sklar** : Soit  $F$  une fonction de répartition de deux variables aléatoires  $X$  et  $Y$  ayant comme fonctions marginales univariées  $F_1$  et  $F_2$ . Alors il **existe** une copule  $C$  telle que :  $F(x_1, x_2) = C(F_1(x_1), F_2(x_2))$ . **Réciproquement**, pour des fonctions

univariées  $F_1, F_2$  et une copule  $C$ , la fonction  $F$  est une fonction de répartition dont les marginales sont  $F_1$  et  $F_2$ . De plus, si les fonctions marginales sont continues, alors la copule  $C$  est **unique**, et donnée par

$$C(u_1, u_2) = F(F_1^{-1}(u_1), F_2^{-1}(u_2)). \quad (3.16)$$

**Définition 3.3.2.2.** *Densité de copule : A partir du théorème précédent et, en dérivant (3.16), la densité d'une copule bivariée est*

$$c(u_1, u_2) = \frac{\partial^2 C}{\partial u_1 \partial u_2} = \frac{f(F_1^{-1}(u_1), F_2^{-1}(u_2))}{f_1(F_1^{-1}(u_1)) f_2(F_2^{-1}(u_2))}, \quad (3.17)$$

et donc

$$f(x_1, x_2) = f_1(x_1) f_2(x_2) c(F_1(x_1), F_2(x_2)) \quad (3.18)$$

où  $x_1 = F_1^{-1}(u_1)$  et  $x_2 = F_2^{-1}(u_2)$ .

Un exemple de copules usuelles est la copule **indépendante** :

$$C(u_1, u_2) = u_1 u_2, \quad (3.19)$$

avec sa densité  $c(u_1, u_2) = 1, (u_1, u_2) \in [0, 1]^2$ .

**Proposition 3.3.2.** *Une copule quelconque  $C(u_1, u_2)$ , satisfait l'inégalité*

$$W(u_1, u_2) \leq C(u_1, u_2) \leq M(u_1, u_2), \quad (3.20)$$

où  $M(u_1, u_2)$  est la **copule de Fréchet-Hoeffding minimale** (ou copule comonotone) :

$$M(u_1, u_2) = \min(u_1, u_2), (u_1, u_2) \in [0, 1]^2, \quad (3.21)$$

et  $W(u_1, u_2)$  est la **copule de Fréchet-Hoeffding maximale** (ou copule contra-monotone) :

$$W(u_1, u_2) = \max\{u_1 + u_2 - 1, 0\}, (u_1, u_2) \in [0, 1]^2. \quad (3.22)$$

Il y a plusieurs familles de copules parmi lesquelles, les **copules archimédiennes** qui forment une des classes les plus importantes ([9] page 109) généralisant les copules usuelles et simple d'utilisation car s'ajustant à plusieurs comportements de dépendance. Elles peuvent s'écrire sous la forme

$$C(u_1, u_2) = \varphi^{-1}(\varphi(u_1) + \varphi(u_2)) \quad (3.23)$$

où  $\varphi(\cdot)$  est une fonction strictement décroissante et convexe telle que  $\varphi(1) = 0$  appelée la génératrice. La fonction de répartition peut aussi s'écrire dans ce cas sous la forme

$$F(x_1, x_2) = \varphi^{-1}(\varphi(F_1(x_1)) + \varphi(F_2(x_2))). \quad (3.24)$$



### 3.3.3 Copule multivariée

Une copule est une fonction de répartition, définie sur  $[0, 1]^n$  dont les marges sont uniformes sur  $[0, 1]$ . Une caractérisation est alors que  $C(u_1, \dots, u_n) = 0$  si une des composantes  $u_i$  est nulle,  $C(1, \dots, 1, u_i, 1, \dots, 1) = u_i$ , et  $C$  est  $n$ -croissante c'est-à-dire que le  $C$ -volume est positif.

Le théorème de Sklar dans le cas multivarié s'énonce de la manière suivante : si  $C$  est une copule, et si  $F_1, \dots, F_n$  sont des fonctions de répartition (univariées), alors  $F(x_1, \dots, x_n) = C(F_1(x_1), \dots, F_n(x_n))$  est une fonction de répartition de dimension  $n$ , dont les marges sont précisément  $F_1, \dots, F_n$ .

Et réciproquement, si  $F$  est une fonction de répartition en dimension  $n$ , il existe une copule  $C$  telle que  $F(x_1, \dots, x_n) = C(F_1(x_1), \dots, F_n(x_n))$ , où les  $F_i$  sont les lois marginales de  $F$ .

Si ces lois marginales sont toutes continues, la copule  $C$  est alors unique, et donnée par la relation  $C(u_1, \dots, u_n) = F(F_1^{-1}(u_1), \dots, F_n^{-1}(u_n))$ . Dans ce cas, on pourra alors parler de "la" copule associée à un vecteur aléatoire  $(X_1, \dots, X_n)$ .

La copule d'un vecteur aléatoire  $(X_1, \dots, X_n)$  est alors la fonction de répartition du vecteur aléatoire  $(F_1(X_1), \dots, F_n(X_n))$ , que l'on notera parfois  $(U_1, \dots, U_n)$ .

### 3.3.4 Différentes méthodes de construction des copules

#### Construction des copules par la méthode d'inversion

C'est une méthode directement basée sur le théorème de Sklar. Étant donnée  $F(x_1, x_2)$  la fonction de répartition conjointe de deux variables aléatoires  $X$  et  $Y$  avec  $F_1(x_1)$  et  $F_2(x_2)$  leurs fonctions marginales, toutes sont supposées continues, la copule correspondante peut être construite en utilisant l'unique transformation inverse (transformation quantile)  $x_1 = F_1^{-1}(u_1)$ ,  $x_2 = F_2^{-1}(u_2)$ , par

$$C(u_1, u_2) = F(F_1^{-1}(u_1), F_2^{-1}(u_2)), \quad (3.25)$$

avec  $u_1, u_2$  uniformes sur  $[0, 1]$ .

#### Exemple :

Considérons la loi conjointe :

$$F_\alpha(x_1, x_2) = [(1 + e^{-x_1} + e^{-x_2} + (1 - \alpha)e^{-x_1 - x_2})^{-1}]^{-1} \quad \text{avec } x_1, x_2 \in \overline{\mathbb{R}}.$$

En prenant les limites quand  $x_1$  et  $x_2$  tendent vers l'infini successivement dans les expressions précédentes conduisent aux marginales

$$F_1(x_1) = F_\alpha(x_1, \infty) = (1 + e^{-x_1})^{-1},$$

et

$$F_2(x_2) = F_\alpha(\infty, x_2) = (1 + e^{-x_2})^{-1}.$$

Les transformations inverses sont

$$x_1 = F_1^{-1}(u_1) = \ln \left( \frac{u_1}{1 - u_1} \right),$$

et

$$x_2 = F_2^{-1}(u_2) = \ln \left( \frac{u_2}{1 - u_2} \right).$$

Et faisant la substitution dans (3.25), on a :

$$\begin{aligned} C(u_1, u_2) &= \left[ 1 + \frac{1 - u_1}{u_1} + \frac{1 - u_2}{u_2} + (1 - \alpha) \frac{(1 - u_1)(1 - u_2)}{u_1 u_2} \right]^{-1}, \\ &= \left[ \frac{1 - \alpha(1 - u_1)(1 - u_2)}{u_1 u_2} \right]^{-1}, \\ &= \frac{u_1 u_2}{1 - \alpha(1 - u_1)(1 - u_2)} \quad \text{où } \alpha \in [-1, 1). \end{aligned} \quad (3.26)$$

La copule de l'équation (3.26) est celle obtenue dans [43] et est appelée la copule de Ali-Mikhail-Haq (AMH).

### Construction des copules à l'aide des lois de distributions classiques

Il y a multiple distributions dont l'exemple le plus connu est la fonction gaussienne. En appliquant le théorème de Sklar, avec des marginales gaussiennes alors on aboutit à une copule gaussienne. De même les copules de Student sont celles dont les marginales sont les distributions de Student.

Les deux exemples précédents de distributions classiques dites elliptiques, font partie des **copules elliptiques**.

### Construction des copules à l'aide d'une fonction génératrice

Cette classe nécessite le choix d'une génératrice  $\varphi$  telle que nous avons définie précédemment forme la classe des **copules archimédiennes**. D'après l'équation (3.23), pour  $\varphi(t) = -\ln(t)$ , la copule indépendante (3.19) appartient à la classe des copules archimédiennes.

**Remarque 2.** *La construction des copules reste un domaine actif malgré les différentes méthodes (algébrique, géométrique, etc) en cours. Par exemple, comme un outil de construction d'une nouvelle copule, la méthode géométrique ne tient compte que de la définition de copule, sans utiliser la fonction de répartition. Et pour qu'enfin la troisième propriété des copules soit valide, les connaissances a priori sur la représentation géométrique sont utilisées. Ces informations sont basées sur le support de la fonction, ou la représentation graphique de ses différentes sections (horizontale, verticale ou diagonale). Dans ce manuscrit nous proposerons aussi une autre nouvelle méthode de construction de copule. Dans l'article [5] aucune démonstration du théorème qui porte actuellement le nom de A. Sklar fut faite, mais il y a une énumération de cinq théorèmes, dont les trois premiers théorèmes (en omettant le théorème portant sur l'inégalité que doit satisfaire*

### CHAPITRE 3. THÉORIE DES COPULES

---

une copule) d'après A. Sklar réduisent le problème de M. Fréchet [44] au problème de la caractérisation des couples à  $n$  dimensions, la preuve des théorème est donnée dans [45].

### 3.4 Utilisation des copules en statistique

On suppose disposer d'un échantillon  $\{x_1^i, x_2^i\}$  pour  $i = 1, \dots, T$ , et nous voulons choisir la copule correspondante aux variables précédentes. La littérature sur les copules nous donne quelques méthodes statistiques dont nous donnons les principales. Premièrement nous considérons le cas où nous choisissons une densité conjointe paramétrique ayant pour paramètre  $\theta$  et nous cherchons à estimer les paramètres. Il y a les méthodes d'estimation du maximum de vraisemblance, l'inférence des fonctions marginales et la méthode bayésienne. Il y a aussi le problème de la sélection de modèle où nous voulons choisir entre les copules celle qui collera aux mieux aux échantillons et pour cela différents critères ont été proposés

#### 3.4.1 Estimation au sens du maximum de vraisemblance

L'estimation du maximum de vraisemblance (MV) est une méthode statistique couramment utilisée pour faire de l'inférence sur les paramètres de la distribution de probabilité d'un échantillon donné. La modélisation des données du monde réel par l'estimation du maximum de vraisemblance offre un moyen d'ajuster les paramètres libre  $\theta$  de ce modèle pour fournir un bon ajustement. Communément, on suppose que les données tirées d'une distribution en particulier sont indépendantes et identiquement distribuées (iid) avec des paramètres inconnus. L'hypothèse précédente permet de simplifier le problème car la vraisemblance peut ainsi être écrite comme un produit de  $n$  densités de probabilité univariée. À partir de l'équation (3.18), si nous supposons que la copule, les fonctions marginales et les fonctions de densité dépendant du paramètre  $\theta$  (qui pourrait être un vecteur de paramètres).

– La vraisemblance  $L(\theta)$  peut être écrite :

$$L(\theta) = \prod_{i=1}^T c(F_1(x_1^i; \theta), F_2(x_2^i; \theta); \theta) \prod_j^2 f_j(x_j^i; \theta),$$

– et le calcul du logarithme de la vraisemblance  $\mathcal{L}(\theta)$  nous donne :

$$\mathcal{L}(\theta) = \sum_{i=1}^T \log c(F_1(x_1^i; \theta), F_2(x_2^i; \theta); \theta) + \sum_{i=1}^T \sum_{j=1}^2 \log f_j(x_j^i; \theta). \quad (3.27)$$

Cette méthode d'estimation de  $\theta$  en trouvant les valeurs de  $\theta$  qui maximisent  $\mathcal{L}(\theta)$ . Les valeurs de  $\theta$  sont données par l'estimateur :

$$\hat{\theta} = \arg \max_{\theta} \mathcal{L}(\theta). \quad (3.28)$$

#### 3.4.2 Inférence des fonctions marginales (IFM)

Dans un but pour le cas d'implémentation de l'algorithme quand le nombre des paramètres est large, l'IFM est la plus utilisée. Au lieu d'utiliser le MLE pour utiliser en

une seule étape le paramètre  $\theta$ , on utilise la méthode de l'IFM (Pour plus de discussion à propos de cette méthode voir [39] page 299).

On décrit les deux étapes de la manière suivante :

**Première étape** : on estime les paramètres  $\theta_j$  (le plus souvent un vecteur de paramètres) pour chacune des marginales.

La vraisemblance, le logarithme de la fonction de vraisemblance et l'estimateur de  $\theta_j$  sont successivement

$$L_j(\theta_j) = \prod_{i=1}^T f_j(x_j^i; \theta_j), \quad j = 1, 2$$

$$\mathcal{L}_j(\theta_j) = \log L_j(\theta_j) = \sum_{i=1}^T \log f_j(x_j^i; \theta_j).$$

$\hat{\theta}_j = \arg \max_{\theta_j} \mathcal{L}_j(\theta_j) \quad \text{pour } j = 1, 2.$

(3.29)

**Deuxième étape** : Estimer le paramètre  $\theta_c$  de la loi conjointe

$$\mathcal{L}(\theta_c, \hat{\theta}_1, \hat{\theta}_2) = \sum_{i=1}^T \log f(x_1^i, x_2^i; \hat{\theta}_1, \hat{\theta}_2, \theta_c)$$

Alors l'estimateur par la méthode de l'IFM est

$\hat{\theta}_c = \arg \max_{\theta_c} \mathcal{L}(\theta_c, \hat{\theta}_1, \hat{\theta}_2).$

(3.30)

Une implémentation des deux méthodes (IFM et MV) précédentes et une analyse comparative peuvent se faire en utilisant par exemple le langage de programmation et l'environnement mathématique utilisés pour le traitement de données et l'analyse statistique basée sur le logiciel R [46], par exemple si nous considérons deux densités marginales gaussiennes ayant des paramètres connus et d'autre part une densité bivariée gaussienne avec un paramètre connu. A partir des paramètres précédents, nous pouvons ainsi générer des données formant une copule gaussienne. Par la suite, en appliquant l'IFM et le MV séparément pour comparer la validité des résultats obtenus. Différentes autres densités de probabilité dont les paramètres peuvent être explicitement calculés pourraient être choisis pour cette implémentation, et on pourrait dresser un tableau comparatif des deux méthodes.

### 3.4.3 L'approche bayésienne

Comme dans la précédente approche, si nous avons un échantillon de données  $\{x_1^i, x_2^i\}$  pour  $i = 1, \dots, T$  pour lequel une famille de densité de copule paramétrique  $c(\mathbf{x} | \theta)$  et

la fonction de vraisemblance

$$f(\mathbf{x} | \theta) = \prod_{i=1}^T c(F_1(x_1^i | \theta), F_2(x_2^i | \theta)) \prod_{j=1}^2 f_j(x_j^i | \theta),$$

et si nous avons l'information *a priori* concernant le paramètre inconnu  $\theta$  sous la forme d'une densité de probabilité *a priori*  $\pi(\theta)$ , alors l'approche bayésienne consiste à calculer la probabilité *a posteriori*

$$f(\theta | \mathbf{x}) = \frac{\pi(\theta)f(\mathbf{x} | \theta)}{f(\mathbf{x})} = \frac{\pi(\theta)f(\mathbf{x} | \theta)}{\int \pi(\theta)f(\mathbf{x} | \theta) d\theta}$$

et ensuite en choisissant un estimateur pour  $\theta$  à partir de cet *a posteriori*. L'approche globale est de choisir la fonction d'utilité  $u(\theta, \tilde{\theta})$ , de calculer ensuite ses valeurs attendues

$$\bar{u}(\tilde{\theta}) = \int u(\theta, \tilde{\theta})f(\theta | \mathbf{x}) d\theta$$

et de choisir comme point estimateur

$$\hat{\theta} = \arg \min_{\tilde{\theta}} \{ \bar{u}(\tilde{\theta}) \}. \quad (3.31)$$

Il est intéressant d'en déduire que pour différent choix de  $u$  nous avons aussi différents estimateurs classiques pour  $\theta$ .

L'estimation EAP (Expected A Posteriori) :

$$u(\theta, \tilde{\theta}) = \|\theta - \tilde{\theta}\|^2 \longrightarrow \hat{\theta}_{EAP} = \int \theta f(\theta | \mathbf{x}) d\theta.$$

L'estimation MAP (Maximum A Posteriori) :

$$u(\theta, \tilde{\theta}) = \delta(\theta - \tilde{\theta}) \longrightarrow \hat{\theta}_{MAP} = \arg \max_{\theta} f(\theta | \mathbf{x})$$

Il y a un lien entre le MAP et le MV car

$$\hat{\theta} = \arg \max_{\theta} f(\theta | \mathbf{x}) = \arg \max_{\theta} \{ \mathcal{L}(\theta) + \log \pi(\theta) \}. \quad (3.32)$$

avec  $\mathcal{L}(\theta)$  étant la vraisemblance (3.27).

### 3.5 Quelques points et discussion sur les copules

---

Nous avons vu que les copules peuvent donner une grande flexibilité, en ce sens que nous pouvons coller plusieurs fonctions marginales que nous voulons pour les associer aux différentes variables aléatoires, et les fonctions de répartition pourraient aussi varier d'une variable à une autre. Nous pouvons par exemple modéliser une fonction de répartition marginale gaussienne à une variable et une autre fonction de répartition

à la deuxième, et ensuite leur associer n'importe quelle copule de notre choix sur ces fonctions de répartition marginales. Mais une des grandes difficultés, quand on travaille avec les données reste à trouver la "meilleur" ou la "bonne" copule qui collera au mieux à notre donnée du fait qu'il y a un grand nombre de copules qui ont été proposées et étudiées jusqu'aujourd'hui. Le domaine des paramètres pour chaque famille de copule est différente, alors il est très difficile de faire une comparaison parmi ces familles de copules.

L'une des procédures les plus utilisées pour faire la sélection du modèle est de choisir pour la copule qui a la grande vraisemblance. Il y a aussi une comparaison parmi les différentes famille de copules, si nous voyons [47] où il y a eu un calcul de la *distance de Kullback-Leibler* parmi les copules avec des densités  $c_1, c_2$  et  $E_{c_1}$  la valeur d'espérance de  $c_1$  :

$$K(c_1, c_2) = E_{c_1} \left[ \log \left( \frac{c_1(u_1, u_2)}{c_2(u_1, u_2)} \right) \right] = \int_0^1 \int_0^1 \log \left( \frac{c_1(u_1, u_2)}{c_2(u_1, u_2)} \right) c_1(u_1, u_2) du_1 du_2. \quad (3.33)$$

Plus précisément pour des raisons de symétrie la *mesure de la divergence de Jeffreys* [48] définie par

$$J(c_1, c_2) = \int_0^1 \int_0^1 (c_1(u_1, u_2) - c_2(u_1, u_2)) \log \left( \frac{c_1(u_1, u_2)}{c_2(u_1, u_2)} \right) du_1 du_2, \quad (3.34)$$

était utilisée car

$$J(c_1, c_2) = K(c_1, c_2) + K(c_2, c_1).$$

D'autre part si on se réfère à [49] où l'estimation bayésienne pour faire le choix de la famille de copule la plus probable dans une famille d'ensemble de copule a été faite. Le choix de l'information a priori se base sur la mesure du tau ( $\tau$ ) de Kendall, définie pour les variables aléatoires continues  $X$  et  $Y$  comme

$$\tau(X, Y) = 4 \int_0^1 \int_0^1 C(u_1, u_2) dC(u_1, u_2) - 1. \quad (3.35)$$

Par exemple pour la famille des copules archimédiennes ayant pour générateur  $\varphi$  [50] on a

$$\tau(X, Y) = 4 \int_0^1 \frac{\varphi(t)}{\varphi'(t)} dt + 1. \quad (3.36)$$

Le Rho ( $\rho$ ) de Spearman est aussi une autre mesure qui a l'avantage de dépendre uniquement sur la structure de  $C$ , mais ne dépend pas sur les comportements des fonctions de repartition marginales :

$$\rho(X, Y) = 12 \int_0^1 \int_0^1 u_1 u_2 dC(u_1, u_2) - 3 = 12 \int_0^1 \int_0^1 C(u_1, u_2) du_1 du_2 - 3. \quad (3.37)$$

Le tau ( $\tau$ ) de Kendall et le Rho ( $\rho$ ) de Spearman ont un avantage sur le classique Rho de Pearson ou le produit de moment de corrélation qui reflète un degré de *relation linéaire* entre deux variables aléatoires  $X, Y$  avec des espérances  $E(X) = \mu_X, E(Y) = \mu_Y$  et des variances non nulles  $\sigma_X, \sigma_Y$  et qui est défini par

$$r(X, Y) = \frac{\text{cov}(X, Y)}{\sigma_X \sigma_Y} = \frac{E((X - \mu_X)(Y - \mu_Y))}{\sigma_X \sigma_Y},$$

et nous pourrons aussi écrire

$$r(X, Y) = \frac{E(XY) - E(X)E(Y)}{\sqrt{E(X^2) - E^2(X)} \sqrt{E(Y^2) - E^2(Y)}}.$$

Pour n'importe qu'elle fonction croissante  $f$  et  $g$ , nous avons :

$$\rho(X, Y) = \rho(f(X), g(Y)), \quad (3.38)$$

$$\tau(X, Y) = \tau(f(X), g(Y)) \quad (3.39)$$

$$r(X, Y) \neq r(f(X), g(Y)) \quad (3.40)$$

Par exemple, pour la copule Gaussienne bivariée avec une corrélation  $\rho$  on a  $\tau = \frac{2}{\pi} \arcsin \rho$ . Il y a aussi d'autres mesures intéressantes dans le cas de la théorie copule des valeurs extrêmes [39], qui peuvent aider à faire une comparaison parmi différentes famille de copules, pour la *mesure de dépendance de queue* qu'on peut définir de la manière suivante :

**Définition 3.5.0.1.** Soit  $\bar{C}(u_1, u_2) = 1 - u_1 - u_2 + C(u_1, u_2)$  la fonction conjointe de survie de deux variables aléatoires uniformes sur  $(0, 1)$  dont la fonction de repartition jointe est la copule  $C$ . Si  $C$  est telle que  $\lim_{u_1 \rightarrow 1} \frac{\bar{C}(u_1, u_1)}{1 - u_1} = \lambda_U$  existes, alors  $C$  a une dépendance de queue supérieure si  $\lambda_U \in (0, 1]$  et n'a pas de dépendance de queue supérieur si  $\lambda_U = 0$ . De manière similaire, si  $\lim_{u_1 \rightarrow 0} \frac{C(u_1, u_1)}{u_1} = \lambda_L$  existe, alors  $C$  a une dépendance inférieure si  $\lambda_L \in (0, 1]$  et n'a pas une dépendance de queue inférieure si  $\lambda_L = 0$ .

Une discussion et des moyens pratique et techniques pour faire un bon choix de copules peut aussi se trouver dans [51]. D'autres moyens pour faire un choix convenable de modèle de dépendance qui se basent sur l'utilisation de la famille exponentielle a été proposée dans [52].

Un autre aspect qui pourrait être pédagogique et fourni par les copules peut s'illustrer de la manière suivante. Soient  $X$  et  $Y$  deux variables aléatoires ayant pour coefficient de corrélation  $\rho$ . Étant données  $f_1(x_1)$ ,  $f_2(x_2)$  leurs densités de probabilité associées, comment alors peut-on construire de manière simple une densité jointe  $f(x_1, x_2)$  qui ne soit pas une densité gaussienne ?

En utilisant le théorème de Sklar, c'est une tâche très simple de faire la construction de  $f(x_1, x_2)$ . Sans nuire à la généralité, nous considérons par exemple, le cas où les moyennes  $\mu_{x_2} = \mu_{x_1} = 0$ , les variances  $\sigma_{x_1}^2 = \sigma_{x_2}^2 = 1$  et la corrélation  $\rho$ . Et avec

$$f_1(x_1) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}x_1^2\right) \quad \text{avec la fonction de répartition} \quad \Phi_1(x_1) = F_1(x_1) = \frac{1}{2}(1 + \operatorname{erf}\left(\frac{x_1}{\sqrt{2}}\right)),$$

$$f_2(x_2) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}x_2^2\right) \quad \text{avec la fonction de répartition} \quad \Phi_2(x_2) = F_2(x_2) = \frac{1}{2}(1 + \operatorname{erf}\left(\frac{x_2}{\sqrt{2}}\right)).$$

Si nous choisissons n'importe quelle densité de copule existante dans la littérature statistique sur les copules, sauf bien sur on ne prenant pas le cas où  $\rho \neq -1, 0, 1$ ; et si nous notons  $A = (1 - \rho^2)^{-1/2}$  alors

$$c_\rho(u_1, u_2) = A \exp\left(\frac{[\Phi_1^{-1}(u_1)]^2 + [\Phi_2^{-1}(u_2)]^2}{2}\right) \exp\left(\frac{2\rho\Phi_1^{-1}(u_1)\Phi_2^{-1}(u_2) - [\Phi_1^{-1}(u_1)]^2 - [\Phi_2^{-1}(u_2)]^2}{2(1 - \rho^2)}\right),$$



$$c_\rho(u_1, u_2) = A \exp \left\{ \frac{-A^2}{2} \left( (\rho\Phi_1^{-1}(u_1))^2 - 2\rho\Phi_1^{-1}(u_1)\Phi_2^{-1}(u_2) + (\rho\Phi_2^{-1}(u_2))^2 \right) \right\}.$$

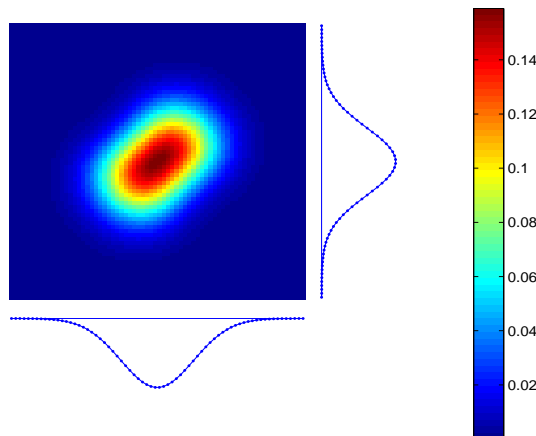
nous obtenons ainsi une densité de probabilité différente de la densité gaussienne. Quand nous choisissons, par exemple la copule de la famille de Farlie-Gumbel-Morgenstern (FGM) :

$$c_\alpha(F_1(x_1), F_2(x_2)) = 1 + \alpha(1 - 2F_1(x_1))(1 - 2F_2(x_2)).$$

La densité de la loi non-Gaussienne est obtenue de la manière suivante :

$$\begin{aligned} f(x_1, x_2) &= f_1(x_1)f_2(x_2)c_\alpha(F_1(x_1), F_2(x_2)) \\ f(x_1, x_2) &= \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}x_1^2\right) \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}x_2^2\right) \left(1 + \alpha \left(1 - 2\frac{1}{2}(1 + \operatorname{erf}\left(\frac{x_1}{\sqrt{2}}\right))\right) \left(1 - 2\frac{1}{2}(1 + \operatorname{erf}\left(\frac{x_2}{\sqrt{2}}\right))\right)\right) \\ &= \frac{1}{2\pi} \exp\left(-\frac{1}{2}(x_1^2 + x_2^2)\right) \left(1 + \alpha \left(1 - (1 + \operatorname{erf}\left(\frac{x_1}{\sqrt{2}}\right))\right) \left(1 - (1 + \operatorname{erf}\left(\frac{x_2}{\sqrt{2}}\right))\right)\right) \\ &= \frac{1}{2\pi} \exp\left(-\frac{1}{2}(x_1^2 + x_2^2)\right) \left(1 + \alpha \left(-\operatorname{erf}\left(\frac{x_1}{\sqrt{2}}\right)\right) \left(-\operatorname{erf}\left(\frac{x_2}{\sqrt{2}}\right)\right)\right) \\ &= \frac{1}{2\pi} \exp\left(-\frac{1}{2}(x_1^2 + x_2^2)\right) \left(1 + \alpha \operatorname{erf}\left(\frac{x_1}{\sqrt{2}}\right) \operatorname{erf}\left(\frac{x_2}{\sqrt{2}}\right)\right) \quad \text{où } \alpha \in [-1, 1] \setminus \{0\}. \end{aligned}$$

Nous donnons la figure suivante pour  $\alpha = 1$  d'un contour d'une densité non-gaussienne  $f(x_1, x_2)$  et ses densités marginales gaussiennes  $f_1(x_1)$ ,  $f_2(x_2)$ .



L'exemple précédent illustre l'aspect pédagogique que peut avoir les copules dans la construction d'une densité non-gaussienne ayant des densités marginales gaussiennes. La discussion initiée dans l'article [53] et les différentes réponses apportées dans [54, 55, 56, 57] ainsi que les références contenues démontrent les aspects avantages et aussi les inconvénients sur l'utilisation des copules pour des applications statistiques. Pour terminer cette partie, le lecteur intéressé d'appliquer la théorie des copules dans un domaine quelconque d'application en statistique est vivement invité à se référer à [58].

Pour conclure cette partie, à notre avis la question du choix des copules en présence des données n'est pas encore définitivement close, en effet cette question fut posée dans [53].

Nous pourrions conclure ce chapitre et cette section concernant l'utilisation des copules et les rappels que nous avons donnés sur les quelques méthodes sur la sélection des modèles statistique dans le cadre des copules par cette réponse (voir [54, 59]) qui s'avère être appropriée :

Model selection is a broad question for which a completely satisfying answer does not yet exist, even in the univariate case. The same strategies can be used here as in many other modeling exercise, i.e choices can be guided by model properties and characterisations diagnostics tools, cross-validation, predictive, accuracy , etc. Given that copula modeling is still in a relatively early stage of development, we concur that much remains to be done in this regard . . .

### 3.6 Conclusion

---

Nous avons rappelé les définitions de quelques notions de statistiques (distributions multivariées, distributions marginales, loi conjointe, etc), et nous avons donnée l'expression des lois classiques de probabilité (loi gaussienne, mélange de lois, loi de Student). En présentant l'équivalence entre les problèmes inverses de recherche d'une loi conjointe à partir de ses lois marginales et la reconstruction d'une image à partir de ses projections, on a aboutit à la notion de copule et leur utilisation. Nous avons donné les principales méthodes statistiques qui permettent de faire le choix d'une copule en présence d'un échantillon. Nous avons aussi rappelé les mesures de dépendance ( la divergence de Kullback-Leibler, le Rho de Spearman, le tau de Kendall, etc) qui permettent d'orienter le choix de l'information a priori. On a conclut ce chapitre en montrant l'avantage des copules, par exemple à facilement construire une densité conjointe non-gaussienne à partir des marginales gaussiennes en éclaircissant ainsi le caractère mal posé du problème de recherche d'une loi conjointe. Une référence bibliographique permet de conclure ce chapitre, celle-ci montre les discussions et aussi les précautions à prendre avant l'utilisation des copules.



# 4

## Copules à maximum d'entropie

Comme nous l'avons rappelé précédemment, la notion de copule fut introduite par A. Sklar [5] et celui-ci a résolu le problème étudié par M. Fréchet [44] et beaucoup d'autres auteurs sur la détermination d'une loi de probabilité conjointe quand les lois marginales univariées sont données. Ce chapitre présente la contribution majeure de mon travail de thèse.

Le problème considéré est celui de la reconstruction d'une image en tomographie dans le cas où nous connaissons seulement deux projections (verticale et horizontale). Évidemment ce problème est mal posé au sens qu'il y a une infinité de solutions possibles. En effet, d'une manière plus générale, le problème de la reconstruction d'image à partir d'un nombre fini de projection est un problème mal posé. Nous constatons alors que le problème de reconstruction considéré avec seulement deux projections est un problème extrêmement mal posé. Si on assimile l'image à reconstruire à une distribution bivariée (sous réserve de contraintes de positivité et de normalisation) alors les deux projections horizontale et verticale seront assimilées à ses deux marginales. Nous voyons alors qu'il y a une similitude avec le problème de M. Fréchet concernant la détermination d'une loi de probabilité conjointe à partir des marginales. Même dans ce cas précis, comme nous l'illustrerons plus tard, le problème reste mal posé. D'où il est nécessaire d'apporter d'autres informations a priori par différentes méthodes dites de régularisation. Parmi les méthodes de régularisation nous avons choisi la méthode de maximum d'entropie, qui consiste à maximiser certaines entropies sous contrainte. Ayant obtenu la loi de densité conjointe correspondante, nous avons ensuite cherché à déterminer la copule associée. Et il s'avère que les copules associées pour certaines entropies que nous exposerons avec plus de détails dans ce chapitre appartient à une nouvelle famille. Nous nous attellerons à prouver les grandes lignes de notre démarche.

### 4.1 Problèmes posés

---

#### 4.1.1 Problème direct

Soient  $f_1(x_1)$  et  $f_2(x_2)$  les deux marginales de  $f(x_1, x_2)$  obtenues par son intégration suivant les deux axes vertical  $x_2$  et horizontal  $x_1$ . Connaissant  $f(x_1, x_2)$  obtenir  $f_1(x_1)$  et  $f_2(x_2)$  est immédiate. En effet, pour n'importe quelle densité de probabilité conjointe

$f(x_1, x_2)$ , calculer les marginales :

$$f_1(x_1) = \int f(x_1, x_2) dx_2 \quad (4.1)$$

et

$$f_2(x_2) = \int f(x_1, x_2) dx_1, \quad (4.2)$$

ne pose aucune difficulté particulière. Pour la plupart des expressions des densités connues, on peut même obtenir des expressions explicites. Un exemple typique est la loi normale

$$\begin{aligned} f(x_1, x_2) &= A \exp \left( \frac{-1}{2(1-\rho^2)} \left( \frac{(x_1 - \mu_1)^2}{\sigma_{x_1}^2} + \frac{(x_2 - \mu_2)^2}{\sigma_{x_2}^2} - \frac{2\rho(x_1 - \mu_1)(x_2 - \mu_2)}{\sigma_{x_1}\sigma_{x_2}} \right) \right) \\ &= \frac{1}{(2\pi)^{|\Sigma|^{1/2}}} \exp \left( -\frac{1}{2} (\mathbf{x} - \boldsymbol{\mu})^t \Sigma^{-1} (\mathbf{x} - \boldsymbol{\mu}) \right) \end{aligned} \quad (4.3)$$

que nous pouvons noter par  $f(\mathbf{x}) = \mathcal{N}(\boldsymbol{\mu}, \Sigma)$  avec  $\boldsymbol{\mu} = [\mu_1, \mu_2]^t$ ,  $\mathbf{x} = [x_1, x_2]^t$ ,

$\Sigma = \begin{bmatrix} \sigma_{x_1}^2 & \rho\sigma_{x_1}\sigma_{x_2} \\ \rho\sigma_{x_1}\sigma_{x_2} & \sigma_{x_2}^2 \end{bmatrix}$ ,  $A = \frac{1}{2\pi\sigma_{x_1}\sigma_{x_2}\sqrt{1-\rho^2}}$  et dont nous connaissons les expressions des marginales

$$f_1(x_1) = \mathcal{N}(\mu_1, \sigma_{x_1}^2) = \frac{1}{\sqrt{2\pi}\sigma_{x_1}} \exp \left( -\frac{1}{2} \frac{(x_1 - \mu_1)^2}{\sigma_{x_1}^2} \right), \quad (4.4)$$

$$f_2(x_2) = \mathcal{N}(\mu_2, \sigma_{x_2}^2) = \frac{1}{\sqrt{2\pi}\sigma_{x_2}} \exp \left( -\frac{1}{2} \frac{(x_2 - \mu_2)^2}{\sigma_{x_2}^2} \right). \quad (4.5)$$

Nous pouvons aussi remarquer que

$$f(x_1, x_2) = f_1(x_1)f_2(x_2)c(x_1, x_2) \quad (4.6)$$

où

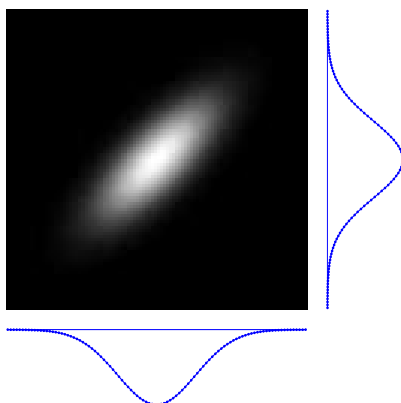
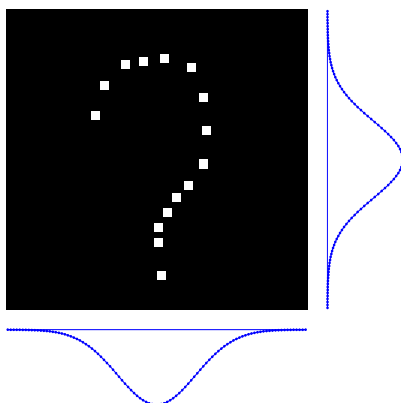
$$\begin{aligned} c(x_1, x_2) &= \frac{f(x_1, x_2)}{f_1(x_1)f_2(x_2)} \quad (4.7) \\ \frac{c(x_1, x_2)}{2\pi A\sigma_{x_1}\sigma_{x_2}} &= \exp \left( -\frac{\rho}{2(1-\rho^2)} \left( \frac{\rho(x_1 - \mu_1)^2}{\sigma_{x_1}^2} + \frac{\rho(x_2 - \mu_2)^2}{\sigma_{x_2}^2} - \frac{2(x_1 - \mu_1)(x_2 - \mu_2)}{\sigma_{x_1}\sigma_{x_2}} \right) \right). \end{aligned}$$

Nous dirons que ce problème est bien posé.

### 4.1.2 Problème inverse

Par contre connaissant seulement  $f_1(x_1)$  et  $f_2(x_2)$ , la recherche de  $f(x_1, x_2)$  est un problème mal posé. En effet, comme nous le savons déjà que toutes les solutions de la forme (4.6) où  $c(u_1, u_2)$  est une densité de copules quelconque sont des solutions possibles satisfaisant les contraintes (4.2). Le choix d'une solution devient alors le choix d'une copule. L'exemple du cas gaussien illustre ce point où  $c(x_1, x_2)$  était une densité de copule gaussienne. Mais sachant que (4.6) est valable pour n'importe quelles autres copules, le choix de la solution dépend du choix arbitraire de copule.

Sans nuire à la généralité, nous considérons par exemple, le cas où les moyennes sont  $\mu_1 = \mu_2 = 0$ , les variances  $\sigma_{x_1}^2 = \sigma_{x_2}^2 = 1$ , et la corrélation  $\rho$ . Pour


 FIGURE 4.1 – Problème direct : connaissant  $f(x_1, x_2)$  trouver  $f_1(x_1)$  et  $f_2(x_2)$ .

 FIGURE 4.2 – Problème inverse : connaissant  $f_1(x_1)$  et  $f_2(x_2)$  déterminer  $f(x_1, x_2)$ .

- $f_1(x_1) = \frac{1}{\sqrt{2\pi}} \exp(-\frac{1}{2}x_1^2)$  avec sa fonction de répartition  $u_1 = \Phi_1(x_1) = F_1(x_1) = \frac{1}{2}(1 + \operatorname{erf}(\frac{x_1}{\sqrt{2}}))$ ,
- $f_2(x_2) = \frac{1}{\sqrt{2\pi}} \exp(-\frac{1}{2}x_2^2)$  avec sa fonction de répartition  $u_2 = \Phi_2(x_2) = F_2(x_2) = \frac{1}{2}(1 + \operatorname{erf}(\frac{x_2}{\sqrt{2}}))$ .

Si nous choisissons n'importe quelle densité de copule existante dans la littérature statistique sur les copules, sauf bien sur en ne prenant pas le cas où  $\rho = -1, 0, 1$ ; et avec  $A = (1 - \rho^2)^{-1/2}$ , nous avons alors

$$c_\rho(u_1, u_2) = A \exp\left(\frac{x_1^2 + x_2^2}{2}\right) \exp\left(\frac{2\rho\Phi_1^{-1}(u_1)\Phi_2^{-1}(u_2) - [\Phi_1^{-1}(u_1)]^2 - [\Phi_2^{-1}(u_2)]^2}{2(1 - \rho^2)}\right),$$

$$c_\rho(u_1, u_2) = A \exp\left\{\frac{-A^2}{2} \left( (\rho\Phi_1^{-1}(u_1))^2 - 2\rho\Phi_1^{-1}(u_1)\Phi_2^{-1}(u_2) + (\rho\Phi_2^{-1}(u_2))^2 \right)\right\}.$$

nous obtenons ainsi une densité de probabilité différente de la densité gaussienne. Quand nous choisissons, par exemple la copule de Farlie-Gumbel-Morgenstern :

$$c_\alpha(F_1(x_1), F_2(x_2)) = 1 + \alpha(1 - 2F_1(x_1))(1 - 2F_2(x_2)) \quad \text{où} \quad \alpha \in [-1, 1] \setminus \{0\}.$$

La densité de la loi non-Gaussienne est obtenue de la manière suivante :

$$\begin{aligned}
 f(x_1, x_2) &= f_1(x_1)f_2(x_2)c_\alpha(F_1(x_1), F_2(x_2)) \\
 f(x_1, x_2) &= f_1(x_1)f_2(x_2) \left( 1 + \alpha \left( 1 - 2\frac{1}{2}(1 + \operatorname{erf}(\frac{x_1}{\sqrt{2}})) \right) \left( 1 - 2\frac{1}{2}(1 + \operatorname{erf}(\frac{x_2}{\sqrt{2}})) \right) \right) \\
 &= \frac{1}{2\pi} \exp\left(-\frac{1}{2}(x_1^2 + x_2^2)\right) \left( 1 + \alpha \left( 1 - (1 + \operatorname{erf}(\frac{x_1}{\sqrt{2}})) \right) \left( 1 - (1 + \operatorname{erf}(\frac{x_2}{\sqrt{2}})) \right) \right) \\
 &= \frac{1}{2\pi} \exp\left(-\frac{1}{2}(x_1^2 + x_2^2)\right) \left( 1 + \alpha \left( -\operatorname{erf}(\frac{x_1}{\sqrt{2}}) \right) \left( -\operatorname{erf}(\frac{x_2}{\sqrt{2}}) \right) \right) \\
 f(x_1, x_2) &= \frac{1}{2\pi} \exp\left(-\frac{1}{2}(x_1^2 + x_2^2)\right) \left( 1 + \alpha \operatorname{erf}\left(\frac{x_1}{\sqrt{2}}\right) \operatorname{erf}\left(\frac{x_2}{\sqrt{2}}\right) \right).
 \end{aligned}$$

Nous donnons sur la figure suivante pour différentes valeurs de  $\alpha$  les contours d'une densité non-Gaussienne  $f(x_1, x_2)$  et ses densités marginales gaussiennes  $f_1(x_1)$ ,  $f_2(x_2)$ .

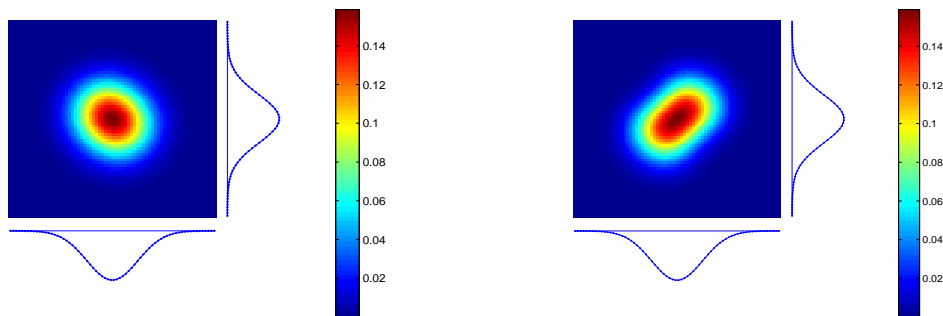


FIGURE 4.3 – Deux solutions différentes pour le problème inverse : connaissant  $f_1(x_1)$  et  $f_2(x_2)$  trouver  $f(x_1, x_2)$ . Les deux densités ont exactement les mêmes marginales.

Il existe différentes manières de régulariser le problème :

## 4.2 Différentes approches pour le problème inverse

### 4.2.1 Fixer a priori une forme pour $f(x_1, x_2)$ .

Par exemple, rechercher  $f(x_1, x_2)$  de la forme séparable

$$f(x_1, x_1) = \lambda_1(x_1)\lambda_2(x_2). \tag{4.8}$$

On cherchera alors les expressions de  $\lambda_1(x_1)$  et  $\lambda_2(x_2)$  en fonction de  $f_1(x_1)$  et  $f_2(x_2)$ . Dans cet exemple une réponse triviale est immédiate.

$$\lambda_1(x_1) = f_1(x_1),$$

$$\lambda_2(x_2) = f_2(x_2).$$

Ou bien généralement une combinaison pour  $a \in \mathbb{R} - \{0\}$  et  $b = 0$

$$\lambda_1(x_1) = af_1(x_1) + b$$

$$\lambda_2(x_2) = \frac{1}{a}f_2(x_2) - b.$$

Un autre exemple est de choisir la structure

$$f(x_1, x_2) = \lambda_1(x_1) + \lambda_2(x_2) - 1. \quad (4.9)$$

Là aussi, il est possible de montrer par certaine combinaison qu'une solution possible est  $\lambda_1(x_1) = f_1(x_1)$  et  $\lambda_2(x_2) = f_2(x_2)$ . Mais, dans cette approche, ce choix de la forme reste arbitraire.

## 4.2.2 Approche ad hoc de construction d'une copule

Une approche ad hoc pour résoudre le problème de la recherche d'une densité de probabilité  $f(x_1, x_2)$  connaissant ses marginales  $f_1(x_1)$  et  $f_2(x_2)$  sans aucune autre contrainte a été formulée dans [60] dont les grandes étapes exposées sont les suivantes :

Trouver une fonction  $f^1(x_1, x_2)$  telle que

$$\int_0^1 \int_0^1 f^1(x_1, x_2) dx_1 dx_2 = 0 \text{ et } \int_0^1 f^1(x_1, x_2) dx_1 = 0, \int_0^1 f^1(x_1, x_2) dx_2 = 0, \quad (4.10)$$

et ensuite calculer pour n'importe quelle fonction  $f$  dont les intégrales définies par

$$\begin{aligned} \int_0^1 \int_0^1 f(x_1, x_2) dx_1 dx_2 &= \Delta \\ \int_0^1 f(x_1, x_2) dx_2 &= f_1(x_1), \text{ et } \int_0^1 f(x_1, x_2) dx_1 = f_2(x_2), \end{aligned} \quad (4.11)$$

existent et sont finies. La méthode dans ([60]) conduit à

$$f(x_1, x_2) = f^1(x_1, x_2) + f_1(x_1) + f_2(x_2) - \Delta. \quad (4.12)$$

Et en utilisant la loi conjointe  $f(x_1, x_2)$  et ses marginales  $f_1(x_1)$  et  $f_2(x_2)$ , l'auteur de [60] obtient une famille de copule de densité ayant la forme :

$$c(x_1, x_2) = 1 + \theta f^1(x_1, x_2) \quad (4.13)$$

avec un paramètre adaptable  $\theta$  qui est introduit si la contrainte de positivité dans l'équation (4.13) n'est pas satisfaite mais que  $f^1$  soit une fonction bornée. D'une manière générale  $1 + \sum_{i=1}^n f_i^1$  est une densité avec  $f_i^1$  remplissant les conditions (4.10). Cette méthode que nous avons décrite permet de générer toutes les copules de formes polynomiales [61] dont nous en parlerons plus tard.



### 4.2.3 Choisir une copule et rajouter d'autres contraintes

Par exemple, si  $f_1(x_1) = \mathcal{N}(\mu_1, \sigma_{x_1}^2)$  et  $f_2(x_2) = \mathcal{N}(\mu_2, \sigma_{x_2}^2)$ , choisissant une copule gaussienne avec une valeur  $\rho$  donnée régularise le problème.

Ici nous avons utilisé deux types d'information : choix d'une copule et une contrainte en fixant  $\rho$ .

On peut se poser la question, de savoir si l'ajout unique de la contrainte  $\rho$  suffirait.

Pour  $f_1(x_1) = \mathcal{N}(\mu_1, \sigma_{x_1}^2)$  et  $f_2(x_2) = \mathcal{N}(\mu_2, \sigma_{x_2}^2)$ , on chercherait la densité  $f(x_1, x_2)$  telle que :

$$\frac{E\{X_1 X_2\} - E\{X_1\}E\{X_2\}}{\{\text{Var}\{X_1\} \text{Var}\{X_2\}\}^{1/2}} = \rho$$

$$\text{Var}(X_i) = \int (x_i - E\{X_i\})^2 f_i(x_i) dx_i \quad \text{où l'espérance } E\{X_i\} = \int x_i f_i(x_i) dx_i.$$

En effet si on dispose des deux marginales gaussiennes, il y a une toujours une infinité de solutions pour  $f(x_1, x_2)$ , la contrainte  $\rho$  supplémentaire introduite permet de fixer la forme de la densité recherchée. Ce cas précis, n'est pas toujours vérifié pour la recherche quelconque d'une densité conjointe à partir de ses marginales.

### 4.2.4 Choisir une fonctionnelle $J(f)$ et l'optimiser

Une troisième approche est de choisir une fonctionnelle  $J(f)$  qui soit convexe (ou concave) et alors rechercher à la maximiser (ou la minimiser) sous les contraintes

$$\left\{ \begin{array}{l} \int_0^1 f(x_1, x_2) dx_2 = f_1(x_1), \quad \forall x_1 \\ \int_0^1 f(x_1, x_2) dx_1 = f_2(x_2), \quad \forall x_2 \\ \int_0^1 \int_0^1 f(x_1, x_2) dx_1 dx_2 = 1, \\ f(x_1, x_2) \geq 0, \end{array} \right. \quad (4.14)$$

c'est l'approche que nous avons choisie, et parmi les fonctionnelles, nous avons choisi

1.  $J_1(f) = - \iint f(x_1, x_2) \ln f(x_1, x_2) dx_1 dx_2;$
2.  $J_2(f) = \frac{1}{1-q} \ln \left( \iint f^q(x_1, x_2) dx_1 dx_2 \right), \quad q \geq 0 \text{ and } q \neq 1;$
3.  $J_3(f) = \iint \ln f(x_1, x_2) dx_1 dx_2;$
4.  $J_4(f) = \frac{1}{q-1} \left( 1 - \iint f^q(x_1, x_2) dx_1 dx_2 \right), \quad q \geq 0 \text{ and } q \neq 1.$

$J_1(f)$  connue comme l'entropie de Shannon [19];  $J_2(f)$  est l'entropie définie par Rényi [21];  $J_3(f)$  est l'entropie de Burg [23]. L'entropie  $J_4(f)$  est l'expression de l'entropie

définie par Havrda et Charvát [22] et popularisée par Tsallis [24], que nous appellerons entropie de Tsallis-Havrda-Charvát ou bien plus simplement l'entropie de Tsallis.

Il y a des liens entre ces différentes expressions d'entropie. L'entropie  $J_1(f)$  peut être obtenue comme cas particulier des entropies  $J_2(f)$  et  $J_4(f)$  quand  $q \rightarrow 1$ .

De même, on a  $\forall q$  en tout point et toutes les fois que  $J_2$  est maximale, ainsi l'est  $J_4$  et vice versa d'après les relations [62] :

$$\begin{aligned} J_2 &= \frac{1}{1-q} \ln [(q-1)J_4 + 1], \\ J_4 &= \frac{1}{q-1} [1 - e^{(1-q)J_2}]. \end{aligned} \quad (4.15)$$

### 4.3 Entropie en mécanique statistique

---

En mécanique statistique, la formule de Boltzmann définit l'entropie microcanonique d'un système physique à l'équilibre macroscopique, mais laissé libre d'évoluer à l'échelle microscopique entre  $\Omega$  micro-états différents, est donnée par :

$$S = k \ln \Omega \quad (4.16)$$

où  $k$  est la constante de Boltzmann (on suppose souvent  $k = 1$  pour la concision des calculs) .  $\Omega$  est le nombre de configurations du système. Cette formule est gravée sur la stèle du monument funéraire de Ludwig Eduard Boltzmann (1844-1906) au cimetière central de Vienne (Autriche), sous la forme  $S = k \log W$ .

L'introduction par Boltzmann de cette "interprétation statistique" de l'entropie marque un tournant majeur dans la compréhension du passage d'une dynamique microscopique "réversible" à une évolution macroscopique "irréversible"

Pour un système classique (i.e., une collection de particules classiques) avec un ensemble discret de micro-états, si  $E_i$  est l'énergie du micro-état " $i$ ", et  $P_i$  est sa probabilité d'apparition durant les fluctuations du système, alors l'entropie du système est :

$$S_{GB} = -k \sum_i P_i \ln P_i \quad (4.17)$$

Ainsi l'entropie est définie sur deux niveaux différents de la description du système donné. La formule (4.17) est appelée la formule d'entropie de Gibbs, d'après Josiah Willard Gibbs (1839-1903).

La référence [63] présente une discussion des deux formules (4.16) et (4.17) dans le cadre de la statistique mécanique.

La formalisation de la *statistique non-extensive* définit l'entropie comme une vaste propriété, c'est-à-dire que sa valeur dépend de la quantité de matière présente.

Dans le cas discret l'entropie de Tsallis est :

$$S_q = k \frac{1 - \sum_i P_i^q}{q-1} \quad (4.18)$$

Le paramètre  $q$  est appelé l'**indice d'entropie** ou le paramètre de non extensivité, et  $k = \frac{2^{(q-1)}}{2^{(q-1)} - 1}$  d'après Havrda et Charvát [22].

$S_q$  est non extensive dans le sens où pour deux systèmes indépendants  $A$  et  $B$  on a :

$$S_q(A + B) = S_q(A) + S_q(B) + (1 - q)S_q(A)S_q(B), \quad (4.19)$$

alors que dans le sens usuel les probabilités de leur occurrence se factorise en produit

$$P_{ij}(A + B) = P_i(A)P_j(B), \quad \forall(i, j)$$

et d'après l'équation (4.17), on a

$$S_{GB}(A + B) = S_{GB}(A) + S_{GB}(B), \quad (4.20)$$

c'est-à-dire l'entropie de Gibbs-Boltzmann (4.17) est additive.

- si  $q < 1$ , c'est la super-extensivité, c'est-à-dire que l'entropie de tout le système est plus grande que la somme de ses parties ou qu'il y a super-additivité.
- si  $q = 1$ , c'est l'extensivité normale, c'est-à-dire l'entropie est additive
- si  $q > 1$ , c'est la sous-extensivité, c'est-à-dire que l'entropie de tout le système est plus petite que la somme de ses parties ou qu'il y a sous-additivité.

L'entropie proposée par Tsallis est une généralisation de l'entropie de Boltzmann-Gibbs (4.17). L'entropie de Tsallis conduit à la mécanique statistique non-extensive, dont les fonctions sont typiques des **lois de puissance**, au lieu des lois exponentielles [24]. Pour avoir une version complète et les démonstrations des relations brièvement présentées ici, voir [64].

### 4.3.1 Méthode proposée

Nous rappelons que le problème de maximisation sous contrainte d'entropie date des années 1930 [10] et que les conditions d'existence d'une solution sont aussi connues [65] (voir aussi [66] et [67]).

La méthode utilisée pour obtenir notre solution est celle qui a une **ressemblance** à celle dite "des multiplicateurs de Lagrange". Il y a une ressemblance du fait qu'ici les contraintes associées ne sont pas en nombre fini, mais les contraintes associées sont continues (équation (4.14)).

Pour cette raison, nous supposons que  $f$  existe avec une entropie finie et que les intégrales considérées convergent dans l'intervalle  $[0, 1]$ . Alors nous définissons les équations suivantes qui ont alors un sens et sont fondamentales :

$$\begin{aligned} \mathcal{L}_i(f) = & J_i(f) + \lambda_0 \left( 1 - \iint f(x_1, x_2) dx_1 dx_2 \right) + \int \lambda_1(x_1) \left( f_1(x_1) - \int f(x_1, x_2) dx_2 \right) dx_1 \\ & + \int \lambda_2(x_2) \left( f_2(x_2) - \int f(x_1, x_2) dx_1 \right) dx_2, \quad \text{pour } i=1, 2, 3, 4 \end{aligned} \quad (4.21)$$

et nous recherchons les points critiques de  $\mathcal{L}_i$  qui sont donnés par le système d'équation :

$$\frac{\partial \mathcal{L}_i(f, \lambda_0, \lambda_1, \lambda_2)}{\partial f} = 0, \quad (4.22)$$

$$\frac{\partial \mathcal{L}_i(f, \lambda_0, \lambda_1, \lambda_2)}{\partial \lambda_j} = 0, \quad \text{pour } j=1, 2, 3. \quad (4.23)$$

La méthode précédente nous permettra de trouver la forme d'une nouvelle famille de lois de probabilité.

En définissant certains cadre de travail et des conditions sur  $f$  (voir les articles [60, 68, 69, 70] et aussi notre Annexe A) d'autres auteurs ont abouti aux familles de lois conjointes nouvelles ou déjà connues. Mais comme nous l'avons déjà montré de manière succincte, la méthode dans [60] n'utilise pas de contrainte d'entropie. Et il y a aussi une méthode dans [68], mais celle-ci n'utilise que l'entropie de Shannon. La méthode dans [68] inspirera aussi [69] et aussi [70] pour la recherche de l'expression d'une loi conjointe. Mais dans tous les travaux précités, la démarche n'est pas intuitive et parfois il n'y a pas d'expression explicite de la copule.

Mais il s'avère que dans notre cas il y a un bien un fondement, qui est l'apparence à la méthode classique de Lagrange et aussi il y a l'utilisation de plusieurs formes d'entropie. Et en plus, nous présentons quand c'est possible les expressions explicites de la copule d'entropie maximale.

Nous nous placerons dans le cadre des **fonctions de répartition à support borné** (voir [71]) et qui prennent des valeurs arbitrairement proches de 0 et 1, et dont leur probabilité  $\mathbb{P}\{0 \leq T \leq 1\} = 1$ . Celles-ci vérifient les conditions :

$$F(0-) = 0, \quad 0 < F(t) < 1 \quad \text{pour } 0 < t < 1, \quad \text{et } F(1) = 1. \quad (4.24)$$

$F(0-)$  signifie la limite « à gauche » de 0 (c'est-à-dire pour le cas  $t < 0$ )

Les fonctions de répartition qui ont leur support contenu dans un intervalle fermé  $[a, b]$  peuvent être traduites dans l'intervalle  $\mathbf{I} = [0, 1]$  à l'aide du changement d'échelle  $t = \frac{x - a}{b - a}$ ,  $x \in [a, b]$ .

### 4.3.2 Résultats principaux

Le théorème suivant peut être considéré alors comme le résultat principal de notre travail

**Théorème 4.3.1.** *Soient  $F_i$ ,  $i = 1, \dots, n$  des fonctions de répartition marginales univariées et absolument continues à support dans  $[0, 1]$  d'une fonction de répartition multidimensionnelle absolument continue  $F$  ayant pour support  $[0, 1]^n$ . Si  $F$  maximise l'entropie de Tsallis (ou l'entropie de Rényi en normalisant) quand l'indice d'entropie est égale à 2, alors  $F$  est donnée par*

$$F(x_1, \dots, x_n) = \sum_{i=1}^n F_i(x_i) \prod_{\substack{j=1 \\ j \neq i}}^n x_j + (1 - n) \prod_{i=1}^n x_i. \quad (4.25)$$

**Preuve 1.** *Nous faisons la démonstration dans le cas bivarié et nous donnons les arguments conduisant au cas multivarié.*

$$f(x_1, x_2) = \exp(-\lambda_1(x_1) - \lambda_2(x_2) - \lambda_0) \quad (\text{entropie de Shannon}) \quad (4.26)$$

$$\frac{f^{q-1}(x_1, x_2)}{\iint_{\mathbf{I}^2} f^q(x_1, x_2) dx_1 dx_2} = \frac{1-q}{q} (\lambda_1(x_1) + \lambda_2(x_2) + \lambda_0) \quad (\text{entropie de Rényi}); \quad (4.27)$$

$$f(x_1, x_2) = (\lambda_1(x_1) + \lambda_2(x_2) + \lambda_0)^{-1} \quad (\text{entropie de Burg}); \quad (4.28)$$

$$f(x_1, x_2) = \frac{1-q}{q} (\lambda_1(x_1) + \lambda_2(x_2) + \lambda_0)^{\frac{1}{q-1}} \quad (\text{entropie de Tsallis-Havrda-Charvát}) \quad (4.29)$$

où  $\lambda_1(x_1)$ ,  $\lambda_2(x_2)$  et  $\lambda_0$  sont obtenus en remplaçant leurs expressions dans (4.14) et en résolvant le système d'équations.

Pour l'entropie de Shannon, on peut obtenir des expressions explicites suivantes :

$$\lambda_1(x_1) = -\ln \left( f_1(x_1) \int_{\mathbf{I}} \lambda_1(x_1) dx_1 \right), \quad (4.30)$$

$$\lambda_2(x_2) = -\ln \left( f_2(x_2) \int_{\mathbf{I}} \lambda_2(x_2) dx_2 \right), \quad (4.31)$$

et

$$\lambda_0 = \ln \left( \int_{\mathbf{I}} \lambda_1(x_1) dx_1 \int_{\mathbf{I}} \lambda_2(x_1) dx_1 \right). \quad (4.32)$$

Et la densité de la loi conjointe est donnée par

$$f(x, y) = f_1(x_1)f_2(x_2). \quad (4.33)$$

Par contre dans les cas des entropies de Rényi, Burg et Tsallis-Havrda-Charvát, en faisant le remplacement dans l'expression (4.14) nous ne pouvons pas obtenir une solution générale explicite pour les paramètres  $\lambda_0$ ,  $\lambda_1(x_1)$ , et  $\lambda_2(x_2)$  comme fonctions explicites de  $f_1(x_1)$  et  $f_2(x_2)$ , et donc les solutions numériques deviennent nécessaires (voir par exemple [72], pour l'entropie de Burg).

## Cas particulier

Mais pour un cas spécial connu aussi comme l'indice de diversité de Simpson [73] c'est-à-dire le cas où l'indice d'entropie de Tsallis-Havrda-Charvát  $q = 2$ , la densité de probabilité est

$$f(x_1, x_2) = -\frac{1}{2} (\lambda_1(x_1) + \lambda_2(x_2) + \lambda_0) \quad (4.34)$$

et nous pouvons obtenir des expressions explicites de  $\lambda_1(x_1)$ ,  $\lambda_2(x_2)$  et  $\lambda_0$  :

$$\lambda_1(x_1) = -2f_1(x_1) + \int_0^1 \lambda_1(x_1) dx_1 + 2, \quad (4.35)$$

$$\lambda_2(x_2) = -2f_2(x_2) + \int_0^1 \lambda_2(x_2) dx_2 + 2, \quad (4.36)$$

$$\lambda_0 = -2 - \int_0^1 \lambda_1(x_1) dx_1 - \int_0^1 \lambda_2(x_2) dx_2. \quad (4.37)$$

En substituant ces expressions dans (4.34) cela nous donne l'expression de la loi de probabilité suivante dans l'intervalle  $[0, 1]$  :

$$f(x_1, x_2) = f_1(x_1) + f_2(x_2) - 1. \quad (4.38)$$

Pour l'entropie de Rényi, dans le cas spécial  $q = 2$  et  $m = \iint_I f^2(x_1, x_2) dx_1 dx_2$ , la densité a la forme

$$f(x_1, x_2) = -\frac{1}{2}m(\lambda_1(x_1) + \lambda_2(x_2) + \lambda_0),$$

et nous pouvons obtenir les expressions explicites de  $\lambda_1(x_1)$ ,  $\lambda_2(x_2)$  et  $\lambda_0$  qui sont :

$$\lambda_1(x_1) = -\frac{2}{m}f_1(x_1) + \int_I \lambda_1(x_1) dx_1 + \frac{2}{m},$$

$$\lambda_2(x_2) = -\frac{2}{m}f_2(x_2) + \int_I \lambda_2(y) dx_2 + \frac{2}{m},$$

et

$$\lambda_0 = -\frac{2}{m} - \int_I \lambda_1(x_1) dx_1 - \int_I \lambda_2(x_2) dx_2.$$

En supposant que  $m = \|f\|_2^2 = \iint_{[0,1]^2} f^2(x_1, x_2) dx_1 dx_2 = 1$ , le résultat obtenu en maximisant l'entropie de Rényi est le même que (4.38) qui donne le résultat dans le cas bivarié. On peut facilement l'étendre dans le cas multivarié sur  $\mathbf{I}^n$  et ce qui donnera alors

$$f(x_1, \dots, x_n) = \sum_{i=1}^n f_i(x_i) - n + 1. \quad (4.39)$$

Et en utilisant les définitions liant une densité de probabilité avec sa fonction de répartition on déduit alors l'expression (4.25). ■

**Remarque 3.** La forme générale de la densité de probabilité peut être obtenue pour n'importe quelle loi de probabilité à support borné en substituant  $x_1$  et  $x_2$  respectivement avec  $\frac{x_1 - x_{\min}}{x_{\max} - x_{\min}}$  et  $\frac{x_2 - y_{\min}}{y_{\max} - y_{\min}}$ . D'autres valeurs d'indice d'entropie  $q \neq 2$  pourraient aussi être utilisées, mais comme nous l'avons déjà expliqué, le choix restreint à  $q = 2$  permet d'aboutir à une forme explicite et montrer la clarté de cette nouvelle méthode de construction de copule que nous proposons.

## 4.4 Famille de copules

Pour le cas de l'entropie de Shannon (4.33)

$$\begin{aligned} F(x_1, x_2) &= \int_0^{x_1} \int_0^{x_2} f(s, t) ds dt \\ &= \int_0^{x_1} \int_0^{x_2} f_1(s) f_2(t) ds dt \\ &= \int_0^{x_1} f_1(s) ds \int_0^{x_2} f_2(t) dt. \end{aligned}$$

Alors la fonction de répartition est

$$F(x_1, x_2) = F_1(x_1) F_2(x_2),$$

et la copule est

$$C(u, v) = F(F_1^{-1}(u), F_2^{-1}(v)) = uv. \quad (4.40)$$

La copule maximale obtenue avec l'entropie de Shannon est la copule indépendante.

Dans l'autre cas particulier ( $q = 2$ ) de l'entropie de Tsallis-Havrda-Charvát (4.38)

$$\begin{aligned} F(x_1, x_2) &= \int_0^{x_1} \int_0^{x_2} f(s, t) ds dt \\ &= \int_0^{x_1} \int_0^{x_2} (f_1(s) + f_2(t) - 1) ds dt \\ F(x_1, x_2) &= x_2 \int_0^{x_1} f_1(s) ds + x_1 \int_0^{x_2} f_2(t) dt - x_1 x_2, \end{aligned}$$

ce qui donne :

$$F(x_1, x_2) = x_2 F_1(x_1) + x_1 F_2(x_2) - x_1 x_2, \quad 0 \leq x_1, x_2 \leq 1 \quad (4.41)$$

et la copule associée est :

$$C(u, v) = u F_2^{-1}(v) + v F_1^{-1}(u) - F_1^{-1}(u) F_2^{-1}(v). \quad (4.42)$$

D'autre part, en combinant le Théorème 4.3.1 et la méthode algébrique d'inversion (voir le corollaire 2.3.7, page 22 de [9]), nous déduisons aussi le corollaire suivant qui donne une nouvelle méthode de reconstruction de famille de copule.

**Corollaire 4.4.1.** Soient  $F_i, i = 1, \dots, n$  choisies avec les propriétés du Théorème 4.3.1, et soit  $F_i^{(-1)}$  les quasi-inverses (voir la définition 2.3.6, page 21 de [9]) de  $F_i$ . Alors pour n'importe quel  $\mathbf{u}$  dans  $[0, 1]^n$ ,

$$C(u_1, \dots, u_n) = \max \left\{ \sum_{i=1}^n u_i \prod_{\substack{j=1 \\ j \neq i}}^n F_j^{(-1)}(u_j) + (1 - n) \prod_{i=1}^n F_i^{(-1)}(u_i), 0 \right\} \quad (4.43)$$

est une copule.

Le théorème 4.3.1 formalise une nouvelle méthode non-intuitive de reconstruction pour déterminer une fonction de répartition multivariée connaissant ses marginales et ensuite le corollaire 4.4.1 que nous avons donné est un outil pour créer des familles de copules. Nous présenterons certains exemples de famille de copule dans le cas bivarié. Il est à préciser que l'indice d'entropie, qui nous a permis d'obtenir ces résultats est  $q = 2$ .

## 4.5 Exemples de famille de copule bivariée

### 4.5.1 Loi Bêta

La densité de la loi Bêta est :

$$f_1(x_1) = \frac{1}{B(a_1, b_1)} x_1^{a_1-1} (1-x_1)^{b_1-1}$$

$$f_2(x_2) = \frac{1}{B(a_2, b_2)} x_2^{a_2-1} (1-x_2)^{b_2-1},$$

où

$$B(a_i, b_j) = \int_0^1 t^{a_i-1} (1-t)^{b_j-1} dt, \quad 0 \leq x_1, x_2 \leq 1 \text{ et } a_i, b_j > 0.$$

Nous considérons seulement, les cas particuliers et intéressants où les paramètres  $a_i$  et  $b_j$  nous permettent d'obtenir l'inverse de la fonction de repartition.

Cas 1 :  $a_i = 1, b_j = 1$  qui correspondent aux lois uniformes  $f_1$  et  $f_2$

$$f_1(x_1) = 1 \rightarrow F_1(x_1) = x_1 \rightarrow F_1^{-1}(u) = u$$

$$f_2(x_2) = 1 \rightarrow F_2(x_2) = x_2 \rightarrow F_2^{-1}(v) = v$$

$$F(x_1, x_2) = x_1 x_2, \quad (4.44)$$

donne la copule indépendante :

$$C(u, v) = uv. \quad (4.45)$$

Cas 2 :  $a_i > 0, b_j = 1$

$$f_1(x_1) = a_1 x_1^{a_1-1} \rightarrow F_1(x_1) = x_1^{a_1} \rightarrow F_1^{-1}(u) = u^{\frac{1}{a_1}}$$

$$f_2(x_2) = a_2 x_2^{a_2-1} \rightarrow F_2(x_2) = x_2^{a_2} \rightarrow F_2^{-1}(v) = v^{\frac{1}{a_2}}$$

$$F(x_1, x_2; a_1, a_2) = x_2 x_1^{a_1} + x_1 x_2^{a_2} - x_1 x_2, \quad (4.46)$$

à partir de (4.59) donne :

$$C(u, v; a_1, a_2) = u v^{\frac{1}{a_2}} + v u^{\frac{1}{a_1}} - u^{\frac{1}{a_1}} v^{\frac{1}{a_2}}, \quad (4.47)$$

qui est une copule définie presque partout dans  $\mathbf{I}$  et des valeurs appropriées de  $a_1, a_2$ .

Si  $a_1 = a_2 = \frac{1}{a}$ , nous avons remarqué que (6.8) peut être écrite sous la forme

$$C(u, v; a) = (uv)^a (u^{1-a} \otimes_1 v^{1-a}), \quad (4.48)$$

où  $a \geq 1$  et  $u \otimes_a v = [u^a + v^a - 1]^{\frac{1}{a}}$  est le produit généralisé [74].



Cas 3 :  $a_i = b_j = 1/2$  qui correspond à la densité la loi arcsinus :

$$\begin{aligned} f_1(x_1) &= \frac{1}{\pi\sqrt{x_1(1-x_1)}} \rightarrow F_1(x_1) = \frac{2}{\pi} \arcsin(\sqrt{x_1}) \rightarrow F_1^{-1}(u) = \sin^2\left(\frac{\pi}{2}u\right) \\ f_2(x_2) &= \frac{1}{\pi\sqrt{x_2(1-x_2)}} \rightarrow F_2(x_2) = \frac{2}{\pi} \arcsin(\sqrt{x_2}) \rightarrow F_2^{-1}(v) = \sin^2\left(\frac{\pi}{2}v\right) \\ F(x_1, x_2) &= \frac{2x_2}{\pi} \arcsin(\sqrt{x_1}) + \frac{2x_1}{\pi} \arcsin(\sqrt{x_2}) - x_1 x_2, \quad 0 \leq x_1, x_2 \leq 1. \end{aligned} \quad (4.49)$$

La copule associée :

$$C(u, v) = u \sin^2\left(\frac{\pi v}{2}\right) + v \sin^2\left(\frac{\pi u}{2}\right) - \sin^2\left(\frac{\pi u}{2}\right) \sin^2\left(\frac{\pi v}{2}\right). \quad (4.50)$$

### 4.5.2 Loi de Kumaraswamy

La loi de Kumaraswamy [75] qui a beaucoup plus d'avantage, et de propriétés que la loi Bêta et aussi facilement manipulable et présenté dans [76] est donnée par

$$f_i(x_i) = a b x_i^{a-1} (1 - x_i^a)^{b-1} \quad (4.51)$$

avec ses fonctions de repartition

$$F_i(x_i) = 1 - (1 - x_i^a)^b, \quad (4.52)$$

et les fonctions quantiles sont données par

$$F_i^{-1}(u_i) = 1 - (1 - u_i^{\frac{1}{b}})^{\frac{1}{a}}, \quad (4.53)$$

où  $a > 0$  et  $b > 0$ .

La copule associée pour les paramètres appropriés  $a$  et  $b$  et ayant les densités de marginales qui sont celles de la loi de Kumaraswamy a la forme :

$$\begin{aligned} C(u_1, u_2) &= u_1 \left(1 - (1 - u_2^{\frac{1}{b}})^{\frac{1}{a}}\right) + u_2 \left(1 - (1 - u_1^{\frac{1}{b}})^{\frac{1}{a}}\right) \\ &\quad - \left(1 - (1 - u_1^{\frac{1}{b}})^{\frac{1}{a}}\right) \left(1 - (1 - u_2^{\frac{1}{b}})^{\frac{1}{a}}\right) \end{aligned} \quad (4.54)$$

Nous remarquons que si  $1/a = 1$  on a l'équation (4.54) qui coïncide avec l'équation (4.48).

### 4.5.3 Loi de puissance bilatère

La densité de la loi de puissance bilatère (traduction de l'anglais "*two-sided power distribution*") [77] est

$$f_i(x_i; a, b) = \begin{cases} b \left(\frac{x_i}{a}\right)^{b-1}, & 0 \leq x_i \leq a \\ b \left(\frac{1-x_i}{1-a}\right)^{b-1}, & a \leq x_i \leq 1 \end{cases} \quad (4.55)$$

où  $0 \leq a \leq 1$ ,  $b > 0$ , et la fonction de répartition

$$F_i(x_i; a, b) = \begin{cases} a \left(\frac{x_i}{a}\right)^b, & 0 \leq x_i \leq a \\ 1 - (1-a) \left(\frac{1-x_i}{1-a}\right)^b, & a \leq x_i \leq 1 \end{cases} \quad (4.56)$$

et

$$F_i^{-1}(u_i; a, b) = \begin{cases} \sqrt[b]{a u_i^{b-1}}, & 0 \leq u_i \leq a \\ 1 - \sqrt[b]{(1-u_i)(1-a)^{b-1}}, & a \leq u_i \leq 1 \end{cases} \quad (4.57)$$

quand  $b = 2$  cela correspond à la loi triangulaire dans le cas continu, et par suite l'inverse de la fonction de répartition est donnée par

$$F_i^{-1}(u_i; a) = \begin{cases} \sqrt{a u_i}, & 0 \leq u_i \leq a \\ 1 - \sqrt{(1-u_i)(1-a)}, & a \leq u_i \leq 1 \end{cases} \quad (4.58)$$

La copule associée pour les paramètres appropriés  $a$  et  $b$  et ayant les densités de marginales qui sont définies par la loi de puissance bilatère a la forme suivante :

Cas 1 :  $0 \leq u_i \leq a$  et  $b > 0$  :

$$C(u_1, u_2) = u_1 \left(\sqrt[b]{a u_2^{b-1}}\right) + u_2 \left(\sqrt[b]{a u_1^{b-1}}\right) - \left(\sqrt[b]{a u_1^{b-1}}\right) \left(\sqrt[b]{a u_2^{b-1}}\right) \quad (4.59)$$

Cas 2 :  $a \leq u_i \leq 1$  et  $b > 0$  :

$$C(u_1, u_2) = u_1 \left(1 - \sqrt[b]{(1-u_2)(1-a)^{b-1}}\right) + u_2 \left(1 - \sqrt[b]{(1-u_1)(1-a)^{b-1}}\right) - \left(1 - \sqrt[b]{(1-u_1)(1-a)^{b-1}}\right) \left(1 - \sqrt[b]{(1-u_2)(1-a)^{b-1}}\right) \quad (4.60)$$

À notre connaissance, les copules définies dans les équations :

(6.8),(4.48),(4.50),(4.54),(4.59),(4.60) sont des nouvelles familles.

#### 4.5.4 Lien avec d'autres familles de copules

Dans le cas bivarié, à partir de (4.43) si on note  $F_1^{-1}(u_1) \equiv \Phi$  et  $F_2^{-1}(u_2) \equiv \psi$ , pour n'importe quelles fonctions  $\Phi$  et  $\psi$ , nous avons

$$C(u, v) = v\Phi(u) + u\psi(v) - \Phi(u)\psi(v) \quad (4.61)$$

avec les propriétés suivantes :

- (1)  $\Phi(0) = \psi(0) = 0$ ,
- (2)  $\Phi(1) = \psi(1) = 1$ .

Par conséquent, pour déterminer les cas pour lesquels  $C$  est une copule, le résultat [78] peut être adapté. Nous pouvons également remarquer que, dans [78], la famille suivante a été proposée :

$$C(u, v) = uv + \theta\Phi(u)\psi(v), \quad \theta \in [-1, 1] \quad (4.62)$$

où  $\Phi$ ,  $\psi$  sont des fonctions de répartition absolument continue sur  $[0, 1]$  et leurs dérivées sont bornées pour presque chaque  $u, v$  dans  $[0, 1]$ , avec les conditions

- (1)  $\Phi(0) = \psi(0) = 0$ ,
- (2)  $\Phi(1) = \psi(1) = 0$ .

La famille dans l'équation (4.62) généralise beaucoup de familles connues de copules mais il est à noter que celle-ci peut ne pas contenir la famille (4.61). En plus, la méthode utilisée pour leur construction est différente. En particulier, si  $\theta = -1$  et  $uv = v\Phi(u) + u\psi(v)$  les deux familles peuvent coïncider. Mais en général, ce n'est pas le cas, il suffit de comparer leurs propriétés statistiques, par exemple en calculant leurs mesures de dépendances, ce qui se fera dans la section suivante.

En résolvant le système des inéquations (voir Théorème 2.2.7 dans [9], page 13) les conditions que nous avons obtenues sur  $\Phi$  et  $\psi$ , telle que (4.61) satisfait toutes les propriétés d'une copule pour tout  $v$  (resp.  $u$ ) dans  $[0, 1]$  telle que la dérivée partielle  $C_u$  (resp.  $C_v$ ) existe, *pour presque* tout  $u$  (resp.  $v$ ), et pour cette valeur de  $u$  et  $v$  nous avons respectivement :

$$\begin{cases} 0 \leq \frac{\partial C(u, v)}{\partial u} \leq 1 \\ 0 \leq \frac{\partial C(u, v)}{\partial v} \leq 1. \end{cases} \quad (4.63)$$

A partir de l'équation (4.61) :

$$\frac{\partial C(u, v)}{\partial u} = \Phi'(u) (v - \psi(v)) + \psi(v). \quad (4.64)$$

De manière similaire

$$\frac{\partial C(u, v)}{\partial v} = \psi'(v) (u - \Phi(u)) + \Phi(u). \quad (4.65)$$

En particulier quand  $\psi(v) = v$  ou  $\Phi(u) = u$ , nous avons la copule indépendante  $C(u, v) = uv$  qui est un membre de la famille de copule (4.61), et les fonctions  $\psi$  ou  $\Phi$  sont bien définies dans  $[0, 1]$ .

En général, il y a quatre différents cas à considérer pour que (4.61) soit bien une copule suivant les choix sur les fonctions  $\psi$  et  $\Phi$  :

Cas 1 : Si  $\psi(v) < v$  (resp. si  $\Phi(u) < u$ ) nous avons respectivement :

$$\begin{cases} 0 \leq \Phi'(u) (v - \psi(v)) + \psi(v) \leq 1 \\ 0 \leq \psi'(v) (u - \Phi(u)) + \Phi(u) \leq 1. \end{cases} \quad (4.66)$$

Alors

$$\begin{cases} 0 \leq \Phi'(u) + \frac{\psi(v)}{v - \psi(v)} \leq \frac{1}{v - \psi(v)} \\ 0 \leq \psi'(v) + \frac{\Phi(u)}{u - \Phi(u)} \leq \frac{1}{u - \Phi(u)}, \end{cases} \quad (4.67)$$

par conséquent

$$\begin{cases} 0 \leq \Phi'(u) + \frac{\psi(v)}{v - \psi(v)} \leq \frac{1}{v - \psi(v)} \\ 0 \leq \psi'(v) + \frac{\Phi(u)}{u - \Phi(u)} \leq \frac{1}{u - \Phi(u)}. \end{cases} \quad (4.68)$$

Finalement, nous avons

$$\begin{cases} 0 \leq \Phi'(u) + \frac{\psi(v)}{v - \psi(v)} \\ 0 \leq \frac{1 - \psi(v)}{v - \psi(v)} - \Phi'(u) \end{cases} \quad (4.69)$$

$$\begin{cases} 0 \leq \psi'(v) + \frac{\Phi(u)}{u - \Phi(u)} \\ 0 \leq \frac{1 - \Phi(u)}{u - \Phi(u)} - \psi'(v). \end{cases} \quad (4.70)$$

Cas 2 : Si  $\psi(v) > v$  (ou resp. si  $\Phi(u) > u$ ) nous avons respectivement :

$$\begin{cases} 0 \leq \Phi'(u) (v - \psi(v)) + \psi(v) \leq 1 \\ 0 \leq \psi'(v) (u - \Phi(u)) + \Phi(u) \leq 1. \end{cases} \quad (4.71)$$

Alors

$$\begin{cases} -\psi(v) \leq \Phi'(u) (v - \psi(v)) \leq 1 - \psi(v) \\ -\Phi(u) \leq \psi'(v) (u - \Phi(u)) \leq 1 - \Phi(u). \end{cases} \quad (4.72)$$

Par cons equent

$$\begin{cases} \frac{1 - \psi(v)}{v - \psi(v)} \leq \Phi'(u) \leq \frac{-\psi(v)}{v - \psi(v)} \\ \frac{1 - \Phi(u)}{u - \Phi(u)} \leq \psi'(v) \leq \frac{-\Phi(u)}{u - \Phi(u)}. \end{cases} \quad (4.73)$$

Finalement, on a

$$\begin{cases} 0 \leq \frac{-\psi(v)}{v - \psi(v)} - \Phi'(u) \\ 0 \leq \Phi'(u) - \frac{1 - \psi(v)}{v - \psi(v)} \\ 0 \leq \frac{-\Phi(u)}{u - \Phi(u)} - \psi'(v) \\ 0 \leq \psi'(v) - \frac{1 - \Phi(u)}{u - \Phi(u)}. \end{cases} \quad (4.74)$$

On note par :

$$A_1(u) = \frac{\Phi(u)}{u - \Phi(u)} \quad , \quad A_2(u) = \frac{1 - \Phi(u)}{u - \Phi(u)}, \quad (4.75)$$

et

$$B_1(v) = \frac{\psi(v)}{v - \psi(v)} \quad , \quad B_2(v) = \frac{1 - \psi(v)}{v - \psi(v)}. \quad (4.76)$$

Le tableau suivant r esume toutes les conditions :

cas 1	cas 2	cas 3	cas 4
$\forall u \text{ if } \Phi(u) < u$	$\forall v \text{ si } \psi(v) < v$	$\forall u \text{ si } \Phi(u) > u$	$\forall v \text{ si } \psi(v) > v$
$0 \leq \psi'(v) + A_1(u)$	$0 \leq \Phi'(u) + B_1(v)$	$0 \leq -\psi'(v) - A_1(u)$	$0 \leq -\Phi'(u) - B_1(v)$
$0 \leq -\psi'(v) + A_2(u)$	$0 \leq -\Phi'(u) + B_2(v)$	$0 \leq \psi'(v) - A_2(u)$	$0 \leq \Phi'(u) - B_2(v)$

**Remarque 4.** *Un autre lien de la nouvelle famille (4.61) peut s' tablir avec d'autres familles de copule. Par exemple un lien avec la famille des copules archim diennes, s' tablirait en calculant le g n rateur  $\varphi$  (d'apr s le Th or me 4.3.8, page 130 du livre [9]) qui est donn e comme solution de l' quation diff rentielle :*

$$\varphi'(u) \frac{\partial C(u, v)}{\partial v} = \varphi'(v) \frac{\partial C(u, v)}{\partial u}. \quad (4.77)$$

## 4.6 Les mesures de dépendances

Pour des variables continues, le tau ( $\tau$ ) de Kendall est :

$$\tau = 4 \int_0^1 \int_0^1 C(u, v) dC(u, v) - 1,$$

le ( $\rho$ ) Spearman est :

$$\rho = 12 \int_0^1 \int_0^1 C(u, v) du dv - 3,$$

et la mesure dite de Blomqvist est :

$$\begin{aligned} \varpi &= 4C\left(\frac{1}{2}, \frac{1}{2}\right) - 1 \\ \varpi &= 2\left(\phi\left(\frac{1}{2}\right) + \psi\left(\frac{1}{2}\right)\right) - 4\phi\left(\frac{1}{2}\right)\psi\left(\frac{1}{2}\right) - 1. \end{aligned} \quad (4.78)$$

Après un ensemble long et fastidieux calculs algébriques simples, leurs expressions respectives sont données dans la ligne suivante en partant de l'équation (4.43) où  $n = 2$  :

$$\tau = 4 \left( \int_0^1 F_1(x_1) dx_1 - 2 \int_0^1 F_1(x_1) dx_1 \int_0^1 F_2(x_2) dx_2 + \int_0^1 F_2(x_2) dx_2 \right) - 2,$$

et

$$\rho = 6 \left( \int_0^1 F_1(x_1) dx_1 - 2 \int_0^1 F_1(x_1) dx_1 \int_0^1 F_2(x_2) dx_2 + \int_0^1 F_2(x_2) dx_2 \right) - 3.$$

Si nous notons par

$$\xi(F_1, F_2) = \int_0^1 F_1(x_1) dx_1 - 2 \int_0^1 F_1(x_1) dx_1 \int_0^1 F_2(x_2) dx_2 + \int_0^1 F_2(x_2) dx_2, \quad (4.79)$$

nous avons la relation entre le tau ( $\tau$ ) de Kendall et le Rho ( $\rho$ ) de Spearman

$$2\xi(F_1, F_2) - 1 = \frac{\tau}{2} = \frac{\rho}{3} \text{ qui implique que } \tau = \frac{2}{3}\rho. \quad (4.80)$$

Nous avons l'expression gamma ( $\gamma_C$ ) de Gini :

$$\begin{aligned} \gamma_C &= 4 \int_0^1 [C(u, u) + C(u, 1 - u)] du - 2 \\ \gamma_C &= 4 \left( \int_0^1 t(g(t) + g(1 - t)) dt + \int_0^1 f(t)(1 - g(t) + g(1 - t)) dt \right) - 2. \end{aligned} \quad (4.81)$$

Et le "Spearman footrule coefficient" ( $\varphi_C$ ) :

$$\begin{aligned} \varphi_C &= 6 \int_0^1 C(u, u) du - 2 \\ \varphi_C &= 6 \left( \int_0^1 t(g(t) + g(1 - t)) dt - \int_0^1 f(t)g(t) dt \right) - 2. \end{aligned} \quad (4.82)$$

Les expressions des mesures de dépendance de la famille de l'équation (4.43) nous montrent que les copules que nous avons introduites n'appartiennent pas à celles étudiées par Rodriguez-Lallena et Úbeda-Flores [78] et qui généralisent de nombreuses familles de copules connues telle que la famille de Fairlie-Gumbel-Morgenstern (FGM).

---

**Remarque 5.** *L'entropie de Shannon, sous les contraintes (4.14) conduit à la copule indépendante.*

*Pour les entropies de Rényi/Tsallis la classe de famille de copules, si elle existe dépendra naturellement de l'indice d'entropie  $q$ . Et les équations (4.27) et (4.29) suggèrent que cette classe est liée aux lois de puissance ("power laws").*

## 4.7 Conclusions

---

En se basant sur deux projections pour reconstruire une image ou bien de deux densités marginales pour retrouver une densité conjointe, le caractère du problème inverse mal posé a été présenté dans le cas de la tomographie et des statistiques. En effet les deux problèmes sont alors équivalents (tomographie et statistique) et une illustration à l'aide d'une densité gaussienne est présentée. Les différentes approches pour résoudre ce problème ont été montrées. Nous avons proposé notre approche qui utilise différents types d'entropie, à savoir les entropies de Shannon, de Rényi et de Tsallis. Et le cas particulier pour l'indice d'entropie  $q = 2$  nous a permis d'obtenir des formes explicites de densités conjointes. Nous avons alors reformulé notre méthode proposée à l'aide d'un théorème permettant de construire une loi multivariée conjointe en ayant la seule connaissance de ses marges univariées sous contrainte des entropies. Et nous en avons déduit les expressions explicites des copules maximisant ces entropies. Des exemples de famille de copules sont données en utilisant les lois Bêta, Kumaraswamy, et les lois puissances bilatères. Nous avons aussi montré un lien avec d'autres familles classiques de copules telles que les copules archimédiennes et les copules de Fairlie-Gumbel-Morgenstern (FGM). Nous avons aussi calculer les expressions des mesures de dépendances (mesure de Blomqvist, tau de Kendall, Rho de Spearman) qui associées à cette famille que nous avons introduite.



# 5

## Copules et tomographie discrète

Dans ce chapitre nous présentons d'abord les notions de copules discrètes, chaînes de Markov, matrices bistochastiques et le problème de tomographie discrète. Ensuite, nous montrons les liens entre les copules discrètes et la tomographie discrète. L'objectif principal est de montrer que ces liens ouvrent une grande perspective pour la résolution des problèmes de reconstruction d'image binaire ou discrète à partir d'un nombre limité de projections.

### 5.1 Introduction

---

La notion de « copule » en statistique est introduite en 1959 [5], avec un intérêt croissant d'applications dans plusieurs domaines qui nécessitent un modèle pour une loi de probabilité jointe ou bien la structure de dépendance entre plusieurs variables aléatoires [39]. Brièvement une copule est une densité de probabilité dont les marginales sont uniformes sur l'intervalle  $[0, 1]$ .

En 1992, William F. Darsow, Bao Gia Nguyen et Elwood T. Olsen ont fait la connexion entre les copules et les processus Markoviens [79] en introduisant une opération binaire sur l'ensemble des copules, appelée « produit de copules ». Mayor *et al* suivi par Kolesárová *et al.* [80, 81] ont donné une version discrète appelée « produit de copules discrètes » qui est reliée à l'ensemble des matrices doublement stochastiques.

La tomographie s'intéresse aux problèmes de reconstruction de la structure interne d'un objet à partir de ses projections (une collection de données prises de l'extérieur de l'objet suivant différents angles de vue). La tomographie a un très large domaine d'application. Comme expliqué dans les chapitres précédents, dans le cas de l'imagerie médicale, la tomographie à rayons X, reconstruit les différentes sections du corps humain à partir d'une collection de données radiographiques prises tout autour et permet ainsi de faire de diagnostic de certaines maladies. Mais aussi dans le cadre industriel, quand on utilise les techniques de contrôle non-destructif (CND), la tomographie aide aussi à détecter les bulles d'air dans une fonte d'acier, la présence ou l'absence d'un défaut dans une construction peut être interprétée par des nombres binaires qui sont souvent "0"="absence " et/ou "1"="présence "[82].

Lorsqu'on utilise les rayons X, le lien entre un point de mesure et l'objet devient une intégration de la fonction qui représente l'objet sur la ligne droite du rayon entre l'émetteur et le détecteur.



Du point de vue mathématique, les objets correspondent à une fonction et le problème posé est alors de reconstruire cette fonction à partir de ses projections. Quand on dispose seulement d'un nombre limité de projections, soit seulement deux projections (horizontale et verticale) nous pouvons voir le lien direct entre la construction d'une loi de probabilité jointe connaissant ses lois marginales à l'aide des copules et la reconstruction d'une image en tomographie. En tomographie analytique le domaine et l'ensemble de toutes les valeurs de la fonction sont supposés continus. Quand un grand nombre de projections uniformément distribuées autour l'objet est disponible, une reconstruction précise peut être faite. Mais en pratique, particulièrement en contrôle non-destructif, les angles de vue sont limités. Le problème de reconstruction d'image devient alors un problème mal posé au sens d'Hadamard. En tomographie discrète le domaine de la fonction peut être aussi bien discret ou continu, et la fonction est à valeurs finies et positives [16].

Brièvement une chaîne de Markov est un processus stochastique ayant la propriété qu'étant donné son état présent, les états futurs sont indépendants des états passés. Ceci signifie que la connaissance de l'état présent suffit pour décrire complètement les informations sur l'évolution future du processus. Dans une chaîne de Markov discrète d'ordre 1, la loi de probabilité de l'état au temps  $t + 1$  peut être obtenue à partir de celle de l'état au temps  $t$  à travers une matrice de transition. Cette matrice de transition a la propriété que la somme de chacune de ses lignes soit égale à 1 (matrice stochastique). Si la chaîne de Markov est réversible, alors sa matrice de transition a aussi la somme de chacune de ses colonnes égale à 1 (Matrice doublement stochastique ou bistochastique). Nous déduirons par la suite le lien entre une matrice bistochastique et les copules discrètes.

Nous rappelons les définitions et les propriétés du produit des copules et le lien avec les chaînes de Markov à travers les copules discrètes. Dans le cas continu, ce fut les objectifs des chapitres précédents où nous avons montré la connection entre les copules et la reconstruction d'image en tomographie et où on s'était basé essentiellement sur l'utilisation de la transformée de Radon.

Ici nous introduisons une nouvelle connection via la notion de copule mais dans **le cas discret** et nous montrons quelques liens avec le problème de la **tomographie discrète**. Pour illustrer notre approche nous choisissons le problème classique de la reconstruction d'une matrice binaire. Et pour conclure, nous faisons une synthèse grâce à un organigramme.

## 5.2 Copule discrète

---

Pour la compréhension des autres parties, nous présentons ici quelques définitions sur les copules discrètes, et ensuite leurs relations avec l'opération matricielle.

### 5.2.1 Produit de copules

Soient  $C = C(u, v)$  une copule et  $D_1C = \frac{\partial C}{\partial u}$ ,  $D_2C = \frac{\partial C}{\partial v}$  ses deux dérivées partielles de premier ordre.

**Définition 5.2.1.1.** Soient  $C_1$  et  $C_2$  deux copules, et définissons le  $*$ -produit de  $C_1$  et  $C_2$  comme :

$$C_1 * C_2 : [0, 1]^2 \longrightarrow [0, 1]$$

$$(u, v) \longmapsto (C_1 * C_2)(u, v) = \int_0^1 D_2 C_1(u, t) \cdot D_1 C_2(t, v) dt$$

**Théorème 5.2.1.** Si  $C_1, C_2$  sont deux copules alors leur produit  $C_1 * C_2$  est une copule.

Les deux premières propriétés des copules sont facilement vérifiables et en appliquant la définition d'un produit de copules on peut aussi montrer que  $(C_1 * C_2)$ -volume du rectangle  $[u_1, u_2] \times [v_1, v_2]$  qui est  $V_{C_1 * C_2}([u_1, u_2] \times [v_1, v_2])$  est positif.

**Remarque** pour des copules  $C, C_1, C_2$ , et  $C_3$

- $\Pi * C = C * \Pi = \Pi$ , ( $\Pi$  est l'élément neutre)
- $M * C = C * M = C$ , ( $M$  est l'élément identité)
- $C_1 * (C_2 * C_3) = (C_1 * C_2) * C_3$  (associativité).

Il s'ensuit que l'ensemble des copules a une structure algébrique de groupe non commutatif, car nous avons par exemple  $(W * C)(x, y) = y - C(1 - x, y)$ , et  $(C * W)(x, y) = x - C(x, 1 - y)$ , c'est-à-dire  $W * C \neq C * W$ . Le produit des copules est une opération binaire analogue au produit matriciel.

## 5.3 Copules décrivant les processus markoviens

---

En théorie de processus stochastique de Markov, l'équation de Chapman-Kolmogorov relie la loi jointe aux différentes lois de probabilité définies sur chacune des composantes du processus.

### 5.3.1 Equations de Chapman-Kolmogorov

Un processus stochastique est une collection  $\{X_t \mid t \in T \subset \mathbb{R}\}$  de variables aléatoires. Désignons respectivement par  $H_{st}$  et  $C_{st}$ , la loi jointe et la copule; ainsi que  $F_s, G_t$ , les lois marginales respectives pour les processus  $X_s$  et  $X_t$ , alors  $H_{st}(x, y) = C_{st}(F_s(x), G_t(y))$ ,  $\forall x, y \in \mathbb{R}$ .  $\{X_t\}$  est un processus de Markov, si pour chaque sous-ensemble fini  $\{t_1, \dots, t_n\}$  de l'ensemble  $T$  et pour  $t \in T$ , sous l'hypothèse que  $\{t_1 < \dots < t_n < t\}$ , la loi conditionnelle dépend seulement de l'état présent du processus, c'est-à-dire

$$P[X_t \leq x \mid X_{i_1} = x_1, \dots, X_{i_n} = x_n] = P[X_t \leq x \mid X_{t_n} = x_n]. \quad (5.1)$$

En utilisant (5.1) et avec  $P(x, s; y, t) = P[X_t \leq y \mid X_s = x]$ , l'équation de Chapman-Kolmogorov est

$$P(x, s; y, t) = \int_{-\infty}^{\infty} P(z, u; y, t) \cdot \frac{\partial P(x, s; z, u)}{\partial z} dz, \quad (s < u < t) \quad (5.2)$$

qui est une marginalisation directement sur la variable de nuisance. Si la loi de probabilité de densité conditionnelle  $p(x, s; y, t) = P[X_t = y \mid X_s = x]$  existe, l'équation de Chapman-Kolmogorov prend la forme

$$p(x, s; y, t) = \int_{-\infty}^{\infty} p(x, s; z, u)p(z, u; y, t) dz, \quad (s < u < t). \quad (5.3)$$

Une chaîne de Markov est donc un processus stochastique ayant la propriété (5.1). Quand la loi de probabilité sur l'espace d'état d'une chaîne de Markov est discrète et que la chaîne est homogène, les équations de Chapman-Kolmogorov peuvent être exprimées en terme matriciel. Si  $X_t$  est l'état du processus au temps  $t$ , alors pour deux points  $i$  et  $j$  de l'espace d'état, nous avons  $p_{ij}(t) = P(X_t = j \mid X_{t_0} = i)$ . Par conséquent la matrice stochastique (ou la matrice de transition) avec des éléments  $p_{ij}$  décriront les transitions d'une chaîne de Markov  $X_t$  sur un espace d'état fini.

### 5.3.2 Équations de Chapman-Kolmogorov via les copules

**Théorème 5.3.1** (Théorème de Darsow-Nguyen-Olsen [79]). *Soit  $\{X_t : t \in T\}$  un processus stochastique.  $\forall s, t \in T$ , soit  $C_{st}$  la copule associée à  $X_s$  et  $X_t$ . Alors les probabilités de transition du processus satisfont les équations de Chapman-Kolmogorov si et seulement si  $C_{st} = C_{su} * C_{ut}$ ,  $\forall s, u, t \in T$  avec  $s < u < t$ .*

Au lieu de la méthode conventionnelle pour la construction d'un processus de Markov, qui utilise une loi de probabilité initiale  $H_{t_0}$  et une famille de loi de probabilités de transition  $P(x, s; y, t)$  vérifiant l'équation de Chapman-Kolmogorov, cette nouvelle alternative avec les copules commence par spécifier toutes les lois marginales et une famille de copule bivariée vérifiant le théorème. En fixant les lois de probabilités de transition d'un processus de Markov et en variant la loi de probabilité initiale varie les lois de probabilités marginales, mais en fixant les copules d'un processus de Markov et en variant la loi de probabilité initiale n'aura aucun effet sur une loi marginale.

## 5.4 Matrices bistochastiques et copule discrète

---

### 5.4.1 Matrices doublement stochastiques

Notant  $\mathfrak{S}_{n \geq 2}$  les groupes de permutation de  $\{1, \dots, n\}$  et  $\mathcal{B} = \{e_1, \dots, e_n\}$  la base canonique de  $\mathbb{R}^n$ .

**Définition 5.4.1.1.** *Si  $\sigma \in \mathfrak{S}_{n \geq 2}$ , la matrice de permutation associée à  $\sigma$  est la matrice  $P_\sigma$  de changement de la base canonique de  $\mathbb{R}^n$  à la base  $\mathcal{B}_\sigma = \{e_{\sigma(1)}, \dots, e_{\sigma(n)}\}$ . Ainsi  $P_\sigma e_j = e_{\sigma(j)}$  alors  $P_\sigma = (\delta_{i, \sigma(j)})_{1 \leq i, j \leq n}$  où  $\delta_{ij} = 1$  si  $i = j$  et  $\delta_{ij} = 0$  si  $i \neq j$ .*

Autrement dit, une matrice de permutation est une matrice obtenue en permutant les colonnes d'une matrice identité  $n \times n$  suivant une permutation des nombres 1 à  $n$ . Chaque ligne et colonne possèdent exactement un seul 1 avec des 0 ailleurs, et chaque permutation correspond à une matrice de permutation unique.

**Définition 5.4.1.2.** Une matrice doublement stochastique (ou bistochastique) est une matrice  $A = (a_{ij})_{1 \leq i, j \leq n}$  telle que  $\forall k \in \{1, \dots, n\}$ ,  $a_{ij} \geq 0$  et  $\sum_{k=1}^n a_{ik} = \sum_{k=1}^n a_{kj} = 1$ .

Ainsi, une matrice doublement stochastique est à la fois stochastique à gauche et à droite.

**Théorème 5.4.1** (Théorème de Birkhoff-Von Neumann [83]). Une matrice  $A = (a_{ij})_{1 \leq i, j \leq n}$  est une matrice bistochastique si et seulement si elle est une combinaison linéaire convexe de matrices de permutation c'est-à-dire  $A = \sum_{k=1}^p \lambda_k P_k$ , où  $P_k$  sont des matrices de permutation et  $\sum_{k=1}^p \lambda_k = 1$ ,  $\lambda_k \in \mathbb{R}_+$ , et  $p \leq (n-1)^2 + 1$ .

## 5.4.2 Copule Discrète

**Définition 5.4.2.1.** Soit  $\mathcal{I}_n = \{0, \frac{1}{n}, \frac{2}{n}, \dots, \frac{n-1}{n}, 1\}$ . Une fonction  $C_{n,m} : \mathcal{I}_n \times \mathcal{I}_m \rightarrow [0, 1]$  est appelée copule discrète sur  $\mathcal{I}_n \times \mathcal{I}_m$  si elle satisfait les conditions suivantes :  $\forall i, k \in \{0, \dots, n\}$ ,  $\forall j, l \in \{0, \dots, m\}$

- $C_{n,m}(\frac{i}{n}, 0) = C_{n,m}(0, \frac{j}{m}) = 0$ ,
- $C_{n,m}(\frac{i}{n}, 1) = \frac{i}{n}$  et  $C_{n,m}(1, \frac{j}{m}) = \frac{j}{m}$ ,
- $C_{n,m}(\frac{k}{n}, \frac{l}{m}) + C_{n,m}(\frac{k-1}{n}, \frac{l-1}{m}) \geq C_{n,m}(\frac{k-1}{n}, \frac{l}{m}) + C_{n,m}(\frac{k}{n}, \frac{l-1}{m})$ .

En considérant le cas particulier où  $n = m$  et avec  $C_{n,n} = C_n$ , la relation entre une copule discrète et une matrice bistochastique est possible grâce à la proposition suivante.

**Proposition 5.4.2.** Pour une fonction  $C_n : \mathcal{I}_n \times \mathcal{I}_n \rightarrow [0, 1]$ , on a l'équivalence :

- $C_n$  est une copule discrète ;
- il existe une matrice bistochastique  $A = (a_{ij})_{1 \leq i, j \leq n}$  telle que  $\forall i, j \in \{0, \dots, n\}$ ,

$$c_{i,j}^{(n)} = C_n\left(\frac{i}{n}, \frac{j}{n}\right) = \frac{1}{n} \sum_{k=1}^i \sum_{m=1}^j a_{km} \quad (5.4)$$

La copule discrète  $C_n$  pourraient être réécrite en terme d'une matrice  $n \times n$  ayant ses éléments donnés par (5.4) c'est-à-dire  $C_n = (c_{i,j}^{(n)})$ . Les éléments d'une matrice bistochastique  $A = (a_{i,j})_{1 \leq i, j \leq n}$  sont donnés par la relation :

$$a_{i,j} = n \left( c_{i+1,j+1}^{(n)} - c_{i,j+1}^{(n)} - c_{i+1,j}^{(n)} + c_{i,j}^{(n)} \right). \quad (5.5)$$

**Définition 5.4.2.2.** Soient  $C_n^1$  et  $C_n^2$  deux copules discrètes définies sur  $\mathcal{I}_n \times \mathcal{I}_n$  et soient respectivement  $B_1$  et  $B_2$  leurs  $n \times n$  matrices bistochastiques correspondantes. Alors la copule discrète  $C_n$  associée à la matrice bistochastique  $A = B_1 \cdot B_2$  est le produit de  $C_n^1$  et  $C_n^2$ . Ce produit est noté par  $C_n = C_n^1 \star C_n^2$ .

En utilisant le théorème de Birkhoff-Von Neumann, il est possible de trouver la matrice  $\{0, 1\}$  – bistochastique associée à une copule [80, 81].

**Remarque 6.** Une question intéressante serait de savoir comment trouver une matrice quelconque  $M$  qui serait la plus proche d'une matrice bistochastique  $A$ , ou de manière équivalente une copule discrète la plus proche à une matrice réelle quelconque  $M$  donnée. La réponse à cette question pourrait permettre alors d'utiliser le lien que nous avons montré précédemment pour toute matrice  $M$ .

Plusieurs techniques d'optimisation sous certaines contraintes et des travaux ont déjà été développés pour résoudre cette question, et se basent essentiellement sur l'utilisation de la norme de Frobenius<sup>1</sup>.

Un travail contenu dans l'article [84] permet de trouver la matrice la plus proche  $A$  à  $M$  si elle existe sous les contraintes que le premier moment, ou bien simultanément le premier et le second moment de  $A$  et  $M$  soient égaux.

Cette technique est l'une des pistes à explorer pour le cas général pour l'utilisation d'une matrice quelconque et ensuite trouver sa matrice bistochastique, et par conséquent aboutir à sa copule discrète associée.

## 5.5 Tomographie discrète

---

Dans cette partie, nous présentons le problème classique en tomographie discrète, qui consiste à reconstruire des matrices binaires à partir de deux projections correspondant à la somme de ses lignes et de ses colonnes. Ryser a trouvé une condition nécessaire et suffisante pour qu'une paire de vecteurs soient les deux projections orthogonales d'un ensemble discret [14]. Dans la preuve de son théorème, il a aussi décrit le premier algorithme de reconstruction pour le cas général d'un ensemble discret à partir des projections orthogonales [15, 85]. Nous rappelons les définitions et les conditions de consistance quand on résout le problème de reconstruction d'une image binaire [16], ensuite nous introduisons notre approche qui utilise la notion de copule discrète.

### 5.5.1 Formulation du problème de reconstruction

Une image composée de  $n \times m$  pixels peut être représentée par une matrice  $A = (a_{ij}), 1 \leq i \leq m, 1 \leq j \leq n$ . Un des problèmes de la tomographie discrète est de trouver les éléments  $a_{ij}$  de la matrice  $A$  étant données le vecteur  $R = (r_1, \dots, r_m) \in \mathbb{N}^m$  dont les composantes sont la somme des lignes de  $A$  et le vecteur  $S = (s_1, \dots, s_n) \in \mathbb{N}^n$  dont les composantes sont la somme de colonnes de  $A$ . Le problème de la reconstruction d'une image binaire ou discrète à partir d'un nombre limité de projections conduit à plusieurs solutions. Il est souvent adéquat de réduire la classe de solutions possibles en introduisant des informations a priori [16].

**Définition 5.5.1.1 (Condition de consistance).** Soit  $\mathfrak{A}(R, S)$  la classe de toutes les matrices binaires  $A = (a_{ij})$  Soient les vecteurs  $R$  et  $S$  tels que  $\sum_{j=1}^n a_{ij} = r_i, \sum_{i=1}^m a_{ij} = s_j$  avec  $1 \leq i \leq m$  et  $1 \leq j \leq n$ .

---

1.  $\|X\|_F = \sqrt{\sum_{i=1}^m \sum_{j=1}^n |x_{ij}|^2} = \sqrt{\text{tr}(X^* X)} = \sqrt{\sum_{i=1}^{\min\{m,n\}} \sigma_i^2}$  avec  $X^*$  la matrice conjuguée transposée de  $X$ ,  $\sigma_i$  sont les valeurs singulières de  $X$ , et "tr" la fonction trace d'une matrice.

Dans le cadre général une paire  $(R, S)$  est compatible s'il existe  $m$  et  $n$  deux entiers positifs tels que

$$R \in \mathbb{N}^m, S \in \mathbb{N}^n, n \geq r_i, m \geq s_j \text{ et } \sum_{i=1}^m r_i = \sum_{j=1}^n s_j. \quad (5.6)$$

Les vecteurs cumulatifs suivants construits à partir de  $A$ , peuvent être vues comme les lois marginales discrètes après normalisation,  $\tilde{r}_i = \sum_{k=1}^i r_k$ ,  $\tilde{s}_j = \sum_{k=1}^j s_k$ , pour  $1 \leq i \leq m$  et  $1 \leq j \leq n$ .

## 5.6 Unicité en tomographie discrète

La notion de composantes échangeables (voir page 16, [82]) est l'outil principal qui permet l'analyse du problème d'unicité en tomographie discrète.

**Définition 5.6.0.2** (Composantes échangeables). *Une composante échangeable d'une matrice binaire  $A$  est la transformation des éléments de  $A$  qui change une matrice partitionnée de  $A$ ,  $A_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$  en  $A_2 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$  ou vice versa et laisse invariant tous les autres éléments.*

**Théorème 5.6.1** (Unicité). *Une matrice binaire est unique (suivant les sommes de ses lignes et de ses colonnes) si et seulement si elle n'a pas de composante échangeable.*

### 5.6.1 Méthode proposée

Avec  $m = n$ , et de la même façon que nous avons défini précédemment une paire compatible de vecteurs  $R$  et  $S$ , nous définissons aussi une *paire compatible de copule discrète* à partir de l'équation (5.5) :

$$\sum_{j=1}^n a_{ij} = n \left( c_{i,1}^{(n)} - c_{i+1,1}^{(n)} + c_{i+1,n+1}^{(n)} - c_{i,n+1}^{(n)} \right) = r_i, \quad i = 1, \dots, n \quad (5.7)$$

$$\sum_{i=1}^n a_{ij} = n \left( c_{1,j}^{(n)} - c_{1,j+1}^{(n)} + c_{n+1,j+1}^{(n)} - c_{n+1,j}^{(n)} \right) = s_j, \quad j = 1, \dots, n. \quad (5.8)$$

D'après les équations (5.7), en appliquant une réduction par télescope nous obtenons :

$$\sum_{i=1}^n r_i = n \left( c_{1,1}^{(n)} - c_{1,n+1}^{(n)} - c_{n+1,1}^{(n)} + c_{n+1,n+1}^{(n)} \right) = \sum_{j=1}^n s_j. \quad (5.9)$$

La relation (5.9) suggère l'existence d'un nombre

$$\gamma = n \left( c_{1,1}^{(n)} - c_{1,n+1}^{(n)} - c_{n+1,1}^{(n)} + c_{n+1,n+1}^{(n)} \right) \in \mathbb{N}. \quad (5.10)$$

Nous avons observé que le nombre  $\gamma$  pourrait être interprété comme une formule explicite qui donne le nombre total de "1" contenu dans la matrice  $A = (a_{ij})$  quand  $a_{ij} \in \{0, 1\}$ . Ceci est un avantage supplémentaire apporté par l'utilisation des copules dans le cas d'un problème de reconstruction d'image binaire. Et enfin, dans les relations liant une copule discrète et la tomographie discrète, la *condition d'unicité* se déduit par la recherche des matrices partitionnées de la matrice  $A = (a_{ij})$ , dont les formes explicites via les copules discrètes sont  $C_1 = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 1 \end{bmatrix}$  et  $C_2 = \begin{bmatrix} 0 & \frac{1}{2} \\ \frac{1}{2} & 1 \end{bmatrix}$ .  $C_1$  et  $C_2$  sont les matrices partitionnées de  $A = (a_{ij})$  qui remplacent respectivement  $A_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$  et  $A_2 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ , pour permettre d'avoir à utiliser le cadre des copules discrètes en tomographie discrète.

### 5.6.2 Construction d'une copule par la méthode algébrique

L'exemple qui suit présente la construction en utilisant une méthode algébrique de la famille de copule de Plackett [86] qui est une généralisation de la table de contingence  $2 \times 2$  :

$a_{11}$	$a_{12}$
$a_{21}$	$a_{22}$

À partir de la proportion du produit croisé :

$$\beta = \frac{a_{11}a_{22}}{a_{12}a_{21}}, \quad (5.11)$$

qui représente la mesure de dépendance entre les variables de la table de contingence  $2 \times 2$ . Avec les étiquettes suivantes pour chacune des catégories des variables "bas" (en abrégé "b") et "haut" (en abrégé "h"), on a  $a_{11} = ("b", "b")$ ;  $a_{12} = ("b", "h")$ ;  $a_{21} = ("h", "b")$ ;  $a_{22} = ("h", "h")$ . Soit  $H$  la loi de probabilité de deux variables aléatoires continues  $X$  et  $Y$ , ayant respectivement les lois marginales  $F$  et  $G$ . Soient  $(x, y) \in \mathbb{R}^2$  et les événements " $X \leq x$ ", " $X > x$ " correspondent aux catégories pour les variables de colonnes "bas" et "haut", et similairement, " $Y \leq y$ ", " $Y > y$ " sont les événements pour les variables de lignes "bas" et "haut". En substituant  $a_{11}, a_{12}, a_{21}, a_{22}$  respectivement par les lois de probabilités  $F(x) - H(x, y)$ ,  $G(y) - H(x, y)$ ,  $H(x, y)$ , et  $1 - F(x) - G(y) + H(x, y)$ . Alors  $\beta$  sera une fonction de  $x$  et  $y$ . Finalement en utilisant le théorème de Sklar, où  $u = F(x) \in [0, 1]$  et  $v = G(y) \in [0, 1]$ , on peut écrire  $\beta$  en terme d'une copule  $C$ ,

$$\beta = \frac{C(u, v) [1 - u - v + C(u, v)]}{[u - C(u, v)] [v - C(u, v)]}. \quad (5.12)$$

En résolvant l'équation (5.12) avec pour inconnue  $C$  (voir page 10, [9]) la famille de copules de Plackett

$$C_\beta(u, v) = \frac{[1 + (\beta - 1)(u + v)] - [[1 + (\beta - 1)(u + v)]^2 - 4uv\beta(\beta - 1)]^{1/2}}{2(\beta - 1)}$$

forment les lois des probabilités pour lesquelles  $\beta$  est une constante.  $\beta = 1$ , donne  $\Pi$  comme solution unique. Quand  $\beta \rightarrow 0$ , et  $\beta \rightarrow \infty$ ,  $C_\beta$  correspondent respectivement à  $W$  et  $M$ .

Cette famille de Plackett dont nous avons présentée la dérivation fut appliquée dans [87] et le résultat était étendu dans le cas multivarié (voir pages 124-128, [39]).

Dans la méthode que nous avons proposé pour la tomographie discrète, la matrice bistochastique associée  $C_\beta(u, v)$  dans chacun de ces cas peut aussi être écrite à l'aide de (5.5), et les  $a_{ij}$  avec les éléments de l'équation (5.11).

Après avoir montré les possibles liens entre la notion de copules discrètes et celle de la tomographie discrète. La similitude entre qu'il y a avec la méthode algébrique de reconstruction d'une copule. Nous donnons un organigramme de synthèse permettant d'être la base de développement d'un schéma de reconstruction d'image en tomographie discrète utilisant les copules discrètes :

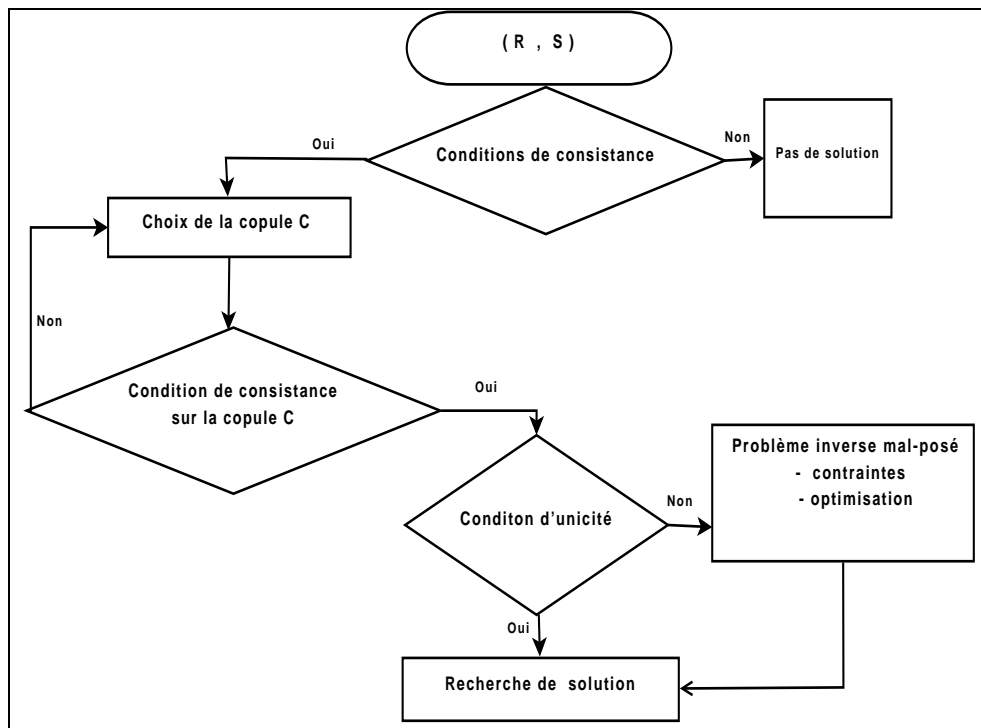


FIGURE 5.1 – Organigramme de synthèse de la méthode proposée

## 5.7 De la copule discrète à la tomographie discrète

L'exemple que nous avons considéré pour expliquer le fondement de la tomographie discrète a été le problème de la reconstitution d'une matrice binaire à partir de la connaissance des sommes de ses éléments suivant les directions horizontale (les lignes de la matrice), verticale (les colonnes de la matrice). Et d'autre part, nous savons aussi qu'une copule permet de **caractériser la dépendance** entre les différentes coordonnées d'une variable aléatoire sans se préoccuper de ses lois marginales. Et nous venons à l'aide du présent chapitre de montrer qu'à chaque copule est associée une matrice bistochastique grâce à l'opération produit de copules. Et nous avons de ce fait proposé une nouvelle méthode en tomographie discrète qui utilise les copules et tout ceci est organisé dans un organigramme.



### 5.7.1 Algorithme pour la matrice bistochastique

Dans l'organigramme précédent, si l'étape de la vérification des conditions de consistance sur  $R$  et  $S$  (voir équation 5.6) est satisfaite alors il faut faire "le choix de la copule". Ensuite il faut vérifier les conditions de consistances sur la copule discrète d'après les équations (5.7). Mais dans notre méthode proposée, il y a un moyen pour trouver la matrice bistochastique  $A$  qui est associée à la matrice de la copule discrète en utilisant l'équation (5.10).

Nous donnons l'algorithme permettant de calculer la matrice bistochastique, en considérant que la copule choisie est la copule de Farlie-Gumbel-Morgenstern (FGM) :

$$C(u, v) = uv + \alpha uv(1 - u)(1 - v). \quad (5.13)$$

En prenant  $\alpha = -1$ , nous avons

$$C(u, v) = uv - uv(1 - u)(1 - v) \quad (5.14)$$

et on fixe la taille  $n$  de la matrice bistochastique.

Le code suivant dont les étapes sont détaillées et qui est écrit en langage MATLAB  $\text{\textcircled{R}}$  permet d'obtenir la matrice bistochastique  $A$  associée à la matrice discrète  $C$  de la copule.

#### Choisir $n$ et initialiser $A$ :

```
%Fixer la valeur de N
N=4;
%Initialiser la matrice A de taille N*N avec des zeros
A=zeros(N);
```

#### Choisir le pas et écrire la copule $C$ :

```
x=[0:1/N:1];
y=[0:1/N:1];

[u,v]=meshgrid(x,y); % creer une grille 2D.

% Ecrire la copule
C=u.*v-u.*v.*(1-u).*(1-v);
```

#### Calculer la matrice $A$ :

En utilisant l'équation l'équation (5.5) pour calculer les éléments de la matrice  $A$ . De manière itérative :

```
for i=1:N
    for j=1:N
```

```

    A(i,j)=N*(C(i+1,j+1)-C(i,j+1)-C(i+1,j)+C(i,j));
end
end

```

On peut aussi calculer A de manière vectorisée :

```
A=N*(C(2:N+1,2:N+1)-C(1:N,2:N+1)-C(2:N+1,1:N)+C(1:N,1:N));
```

## Affichage des matrices C et A :

La fonction *sym* de Matlab permet d'afficher de manière symbolique. Ainsi les commandes suivantes donnent successivement la matrice de la copule discrète C et sa matrice A bistochastique associée.

```

sym(C,'r')
sym(A,'r')

```

On obtient alors la matrice de copule discrète

$$C = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{7}{256} & \frac{5}{64} & \frac{39}{256} & \frac{1}{4} \\ 0 & \frac{5}{64} & \frac{3}{21} & \frac{64}{21} & \frac{1}{2} \\ 0 & \frac{39}{256} & \frac{16}{21} & \frac{135}{256} & \frac{3}{4} \\ 0 & \frac{1}{4} & \frac{1}{2} & \frac{3}{4} & 1 \end{bmatrix},$$

et sa matrice bistochastique

$$A = \begin{bmatrix} \frac{7}{64} & \frac{13}{64} & \frac{19}{64} & \frac{25}{64} \\ \frac{13}{64} & \frac{15}{64} & \frac{17}{64} & \frac{19}{64} \\ \frac{19}{64} & \frac{17}{64} & \frac{15}{64} & \frac{13}{64} \\ \frac{25}{64} & \frac{19}{64} & \frac{13}{64} & \frac{7}{64} \\ \frac{25}{64} & \frac{19}{64} & \frac{13}{64} & \frac{7}{64} \end{bmatrix}.$$

## Sommes des lignes et les colonnes de A :

```

A1=sum(A,1); % somme des lignes de A
A2=sum(A,2); % somme des colonnes de A

```

## Visualiser la matrice bistochastique A :

On utilise la fonction *imagesc* de Matlab.

```

imagesc(A);
colormap(gray);

```

## 5.8 Représentation graphique

La fonction *imagesc(A)* affiche la matrice bistochastique A comme une image. Chaque élément de A correspond à un domaine rectangulaire dans l'image. Les valeurs

des éléments de  $A$  sont celles qui sont indiquées dans *colormap*, ici nous avons choisi “gray”(c’est-à-dire “gris”en français). On notera dans la représentation de graphique de notre image que c’est le choix de  $N$  qui permet de déterminer cette “valeur”qui est en fait l’amplitude lumineuse définissant la couleur ; plus elle est proche du noir, plus la valeur est basse.

En générale pour  $N \neq 2$  les gris, qui sont des teintes du blanc et du noir, sont cependant des couleurs, et non des valeurs. Les “gris ”sont des valeurs particulières sur l’axe noir-blanc. Chacun des gris peut être considéré comme une couleur dépourvue de teinte; le noir et le blanc sont des gris extrêmes. Le noir est un gris de valeur nulle et correspond à l’absence de toute lumière (aucune lumière n’est reçue par l’oeil). Le blanc est un gris de valeur maximale et peut être considéré comme une plénitude de couleurs (l’ensemble des fréquences d’onde lumineuse est reçu par l’oeil avec une valeur maximale).

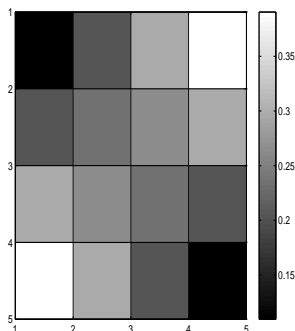


FIGURE 5.2 – Image de la matrice bistochastique avec  $N = 5$

Pour la tomographie discrète binaire, on choisira  $N = 2$  et donc il n’y a que deux valeurs (la blanche et la noire).

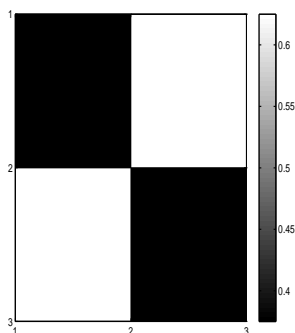


FIGURE 5.3 – Image de la matrice bistochastique dans le cas  $N = 2$

## 5.9 Conclusion

---

Nous avons rappelé le problème de la tomographie discrète. Et le lien entre les copules discrètes et les matrices bistochastiques. En transposant les conditions classiques de consistance et d'unicité utilisée en tomographie dans le cadre des copules discrètes, nous avons proposé une méthode de reconstruction en tomographie discrète qui utilise les matrices bistochastiques. Nous avons aussi montré la méthode algébrique qui permet de reconstruire d'une copule. Il en sort une similitude et un lien avec les méthodes de reconstruction en tomographie discrète. Nous avons alors proposé un organigramme générale et par la suite sur un exemple concret nous avons montré aussi un algorithme qui permet de calculer la matrices bistochastique associée à la matrice d'une copule discrète. Un apport très important de notre travail ici est l'information fournie par l'équation par  $\gamma$  dans l'équation (5.10).



*The mathematical sciences particularly exhibit order, symmetry, and limitation; and these are the greatest forms of the beautiful.*

Aristotle (384 BC-322 BC)



## Expérimentations numériques

Ce chapitre montre les expérimentations numériques à partir des images synthétiques pour illustrer notre nouvelle approche dans le cas de la reconstruction d'un objet en tomographie, quand le nombre de projections est limité à deux. Nous présentons les résultats obtenus en utilisant les copules. Pour illustrer cette nouvelle méthode de reconstruction, nous avons développé un outil de visualisation des copules “**Copula-Density**” et un outil de l'application en tomographie “**Copula-Tomography**” avec une interface graphique dont les codes sources sont écrits en langage MATLAB®.

Nous donnons aussi une illustration d'une de nos familles de copule (équation 4.50), qui contient une belle symétrie et que nous fournissons à titre artistique.

### 6.1 Lien entre la notion de copule et la tomographie

---

À partir de l'équation (2.3) qui donne  $\mathcal{R}f(\theta, r)$  nous considérons le cas particulier où il y a seulement deux projections  $\theta = 0$  et  $\theta = \pi/2$  :

$$\begin{aligned}\mathcal{R}f(0, r) = p_0(r) &= \iint f(x, y)\delta(r - x) \, dx \, dy = \int f(r, y) \, dy \\ \mathcal{R}f(\pi/2, r) = p_{\pi/2}(r) &= \iint f(x, y)\delta(r - y) \, dx \, dy = \int f(x, r) \, dx.\end{aligned}$$

On pose les projections horizontale et verticale respectivement  $f_1(x)$  et  $f_2(y)$

$$f_1(x) = \frac{p_0(x)}{\int p_0(x) \, dx}, \quad f_1(x) \geq 0$$

et

$$f_2(y) = \frac{p_{\pi/2}(y)}{\int p_{\pi/2}(y) \, dy}, \quad f_2(y) \geq 0,$$

permettant de s'assurer que ce sont des densités de probabilité.

Les problèmes inverses de reconstruction et de la détermination de l'image (ou de la fonction de densité)  $f(x, y)$  à partir des projections (ou les lois marginales)  $f_1(x)$  et  $f_2(y)$  sont alors équivalents.

### 6.1.1 En tomographie

Dans le cas des deux projections l'équation (2.7) de **Rétroprojection** devient :

$$f(x, y) = \frac{1}{2}(f_1(x) + f_2(y)). \quad (6.1)$$

De même, l'équation (2.7) de **Rétroprojection filtrée** :

$$f(x, y) = \frac{1}{2} \left( \int \frac{\frac{\partial f_1}{\partial x}(x')}{x' - x} dx' + \int \frac{\frac{\partial f_2}{\partial y}(y')}{y' - y} dy' \right) \quad (6.2)$$

qui peut être aussi implémentée dans le domaine de Fourier :

$$f(x, y) = \int e^{+jux} \text{sign}(u) \left( \int e^{-jux'} f_1(x') dx' \right) du + \int e^{+jvy} \text{sign}(v) \left( \int e^{-jvy'} f_2(y') dy' \right) dv \quad (6.3)$$

$$f(x, y) = \frac{1}{2} \int e^{+jux} |u| \left( \int e^{-jux'} f_1(x') dx' \right) du + \frac{1}{2} \int e^{+jvy} |v| \left( \int e^{-jvy'} f_2(y') dy' \right) dv. \quad (6.4)$$

### 6.1.2 La théorie de copule

La notion de copule nous donne une possibilité pour proposer d'autres méthodes. Alors la première méthode pour le problème de reconstruction que nous proposons est

$$f(x, y) = f_1(x) f_2(y). \quad (6.5)$$

Une deuxième méthode que nous proposons pour la reconstruction des images en tomographie est

$$f(x, y) = f_1(x) f_2(y) c(F_1(x), F_2(y)). \quad (6.6)$$

La fonction matlab *cumsum* permet de calculer  $F_i$ .

En comparant les deux équations (6.1) et (6.5), nous appelons la méthode dans l'équation (6.5) celle de la *Rétroprojection Multiplicative (RPM)*. Et la méthode (6.6) celle de la *Rétroprojection Copule (RPCO)*. On remarque que la méthode de l'équation (6.5) correspondant au cas particulier de la méthode dans l'équation (6.6) où la copule choisie est la copule indépendante de densité égale à l'unité. Le choix de la copule est équivalent à l'introduction d'une information *a priori* sur l'image à reconstruire.

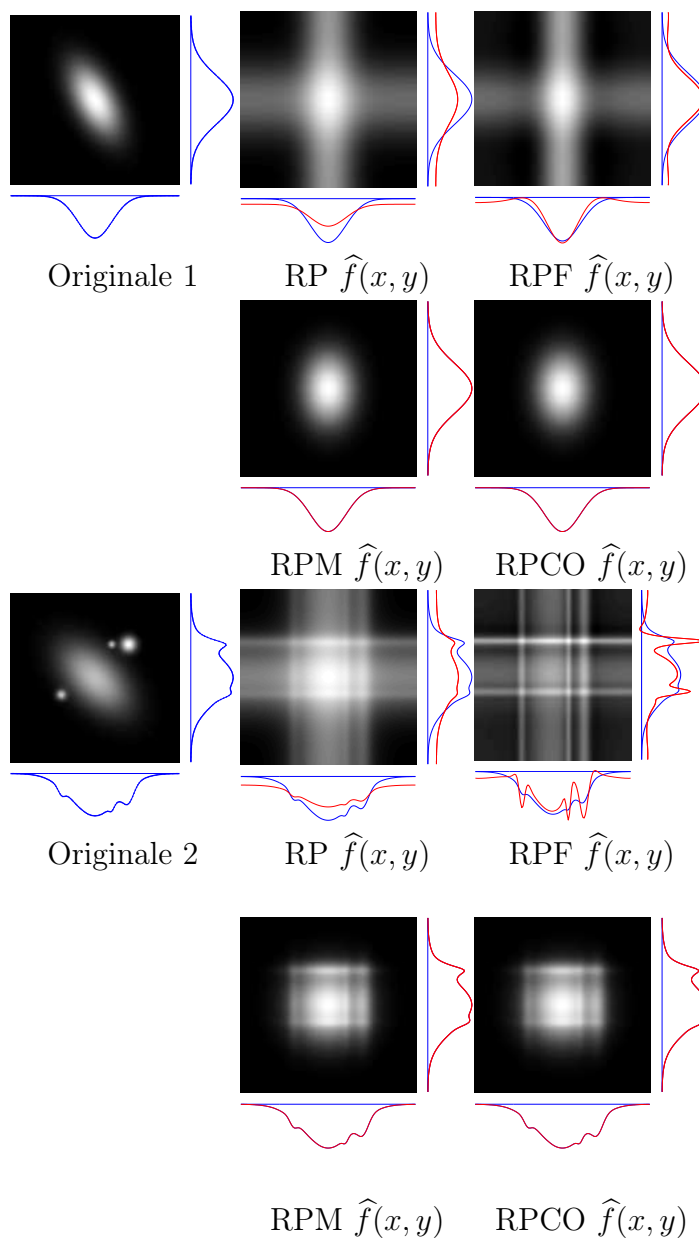


FIGURE 6.1 – Comparaison respective entre RP, RPF, RPM et RPCO sur deux exemples. Sur ces images les courbes en bleu sont les projections des objets originaux et les courbes en rouges sont les projections des objets. On constate que dans RPM et RPCO ces courbes sont superposées.



## 6.2 Quelques représentations graphiques des nouvelles copules

---

Les expressions de deux familles de copules que nous avons obtenues au chapitre 4

$$C(u, v; a_1, a_2) = v u^{a_1} + u v^{a_2} - u^{a_1} v^{a_2}. \quad (6.7)$$

Quand  $a_1 = a_2 = a$ , nous avons vérifié que pour  $1 \leq a \leq 2$ , la famille (6.7) a toutes les propriétés d'une copule. Nous pouvons aussi remarquer ici que les familles (6.8) ont une similitude avec celles qui sont obtenues par la méthode géométrique en [88].

$$C(u, v) = u \sin^2\left(\frac{\pi v}{2}\right) + v \sin^2\left(\frac{\pi u}{2}\right) - \sin^2\left(\frac{\pi u}{2}\right) \sin^2\left(\frac{\pi v}{2}\right). \quad (6.8)$$

Équation (6.8) est une copule bien définie et ne dépend pas d'un paramètre. Nous montrons cette nouvelle copule, ses densités conjointes, et les fonctions de répartition (F.r) dans l'intervalle  $[0, 1]^2$  (voir la figure 6.2). Puis, en remarquant la structure périodique de la fonction sinus dans sa formule et qui est périodique de période 4. Par curiosité artistique nous montrons les tracés de contours [89] de cette copule et de ses densités conjointes sur différents intervalles. Certaines représentations agréables et plaisantes du contour de la nouvelle famille de copule et l'extrapolation en dehors du domaine  $[0, 1]$  en fonction de la structure périodique sont données ci-dessous. On peut noter que ces figures ne sont que pour leurs formes artistiques. Leurs expressions mathématiques ne sont valables que pour l'intervalle  $[0, 1]$  comme sur la figure 6.2.

Figure 6.2 :

Première et deuxième lignes : tracés de contours et les maillages de copule (6.8).

Figure 6.4 :

Gauche : Les intervalles sont  $[0, k], k = \{1, 2, 3, 5, 7, 10\}$

Droite : Les intervalles sont  $[-k, k], k = \{1, 2, 3, 5, 7, 10\}$ .

Figure 6.3 :

Première et deuxième lignes : tracés de contours et les maillages de copule (6.7) pour  $a = 2$ .

**Remarque 7.** *L'équation (6.6) est une contribution importante de notre rapport car cette équation nous montre d'une part le lien entre le problème de reconstruction en tomographie et la notion de copule mais aussi la relation entre la transformée de Radon [1] qui est le fondement mathématique des méthodes de tomographie et le théorème de Sklar [5] qui donne la solution au problème de recherche de la densité conjointe dont les densités marginales sont données.*

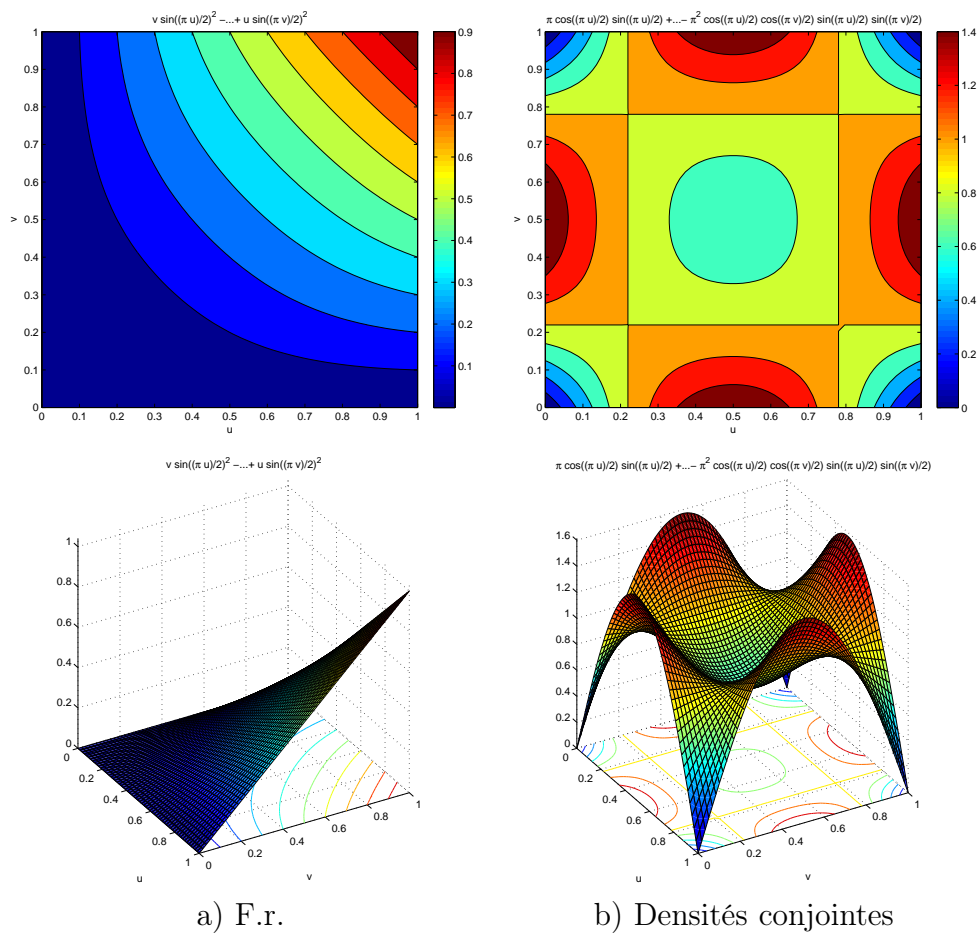


FIGURE 6.2 – Une représentation de la famille (6.8)

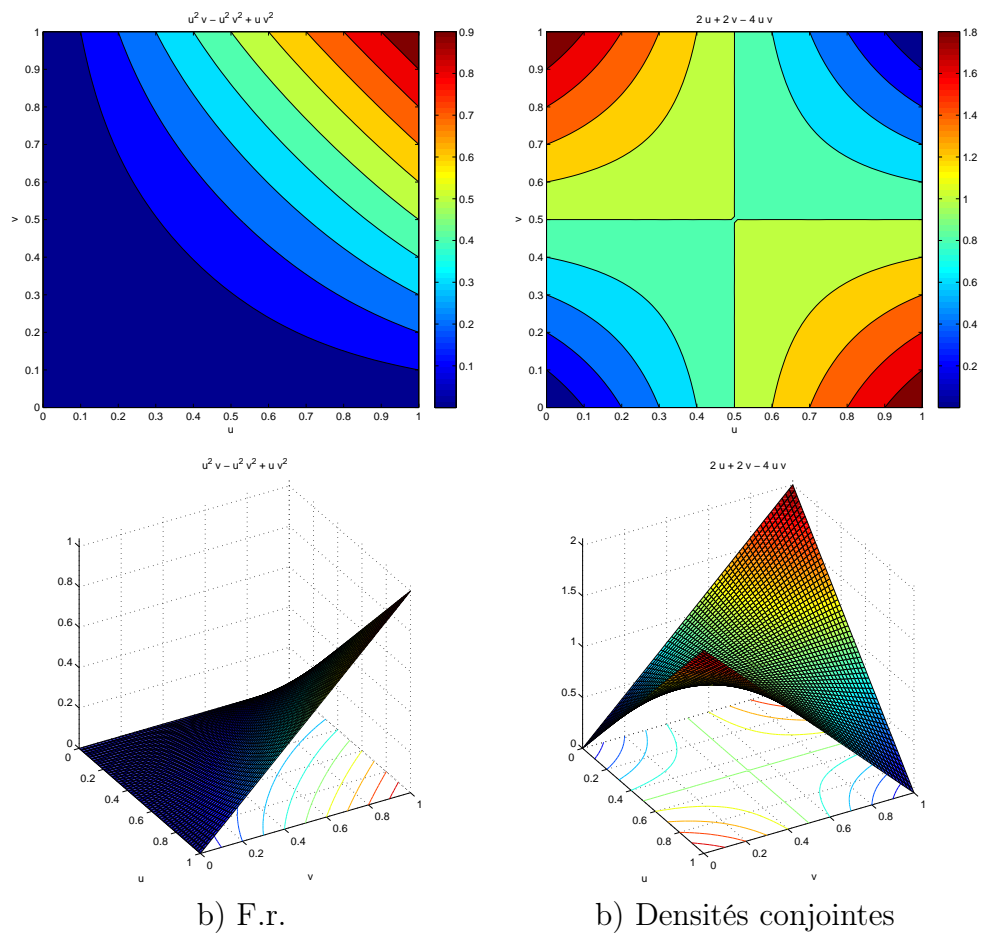


FIGURE 6.3 – Une représentation de la famille (6.7)

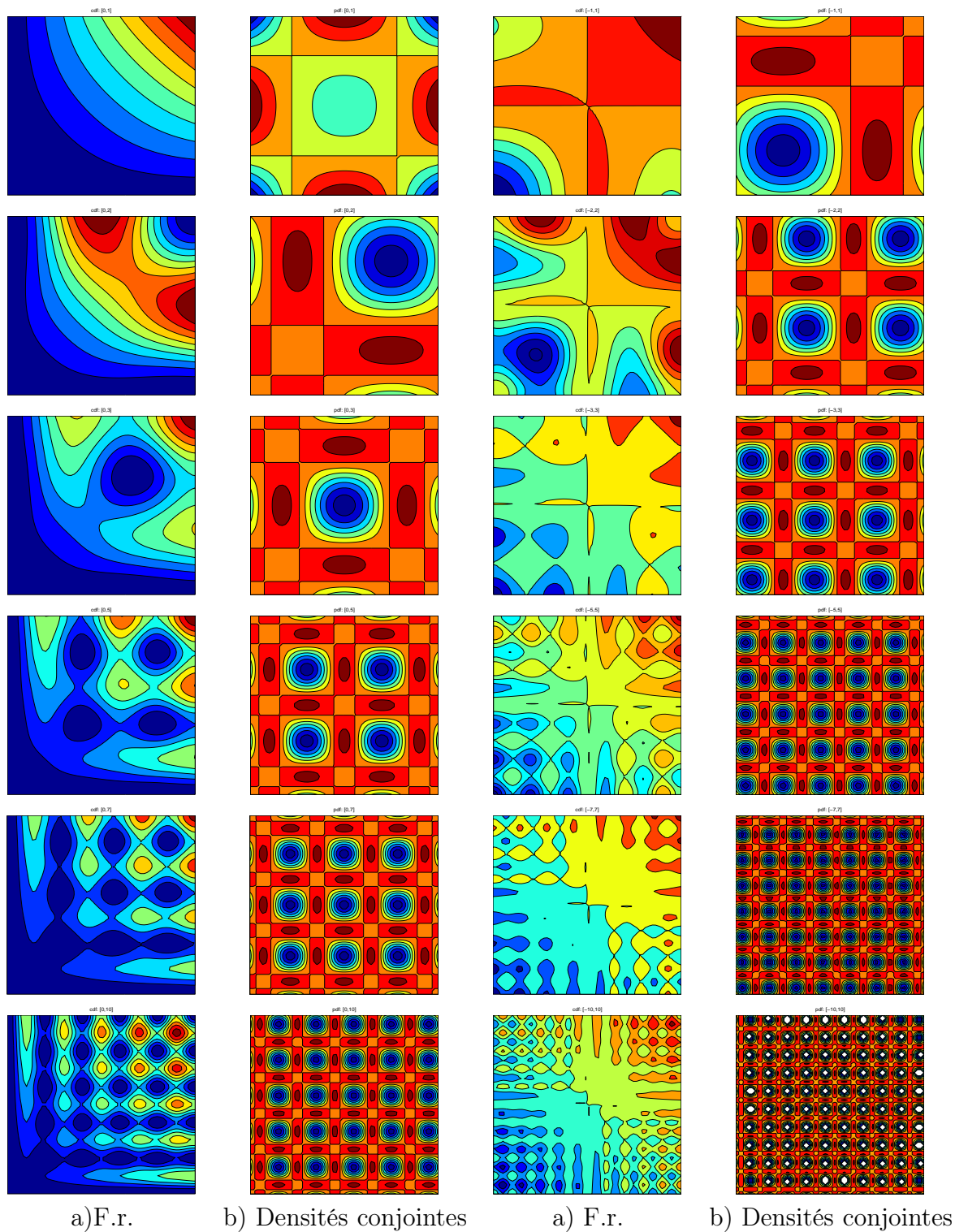


FIGURE 6.4 – Une représentation de la copule (6.8) en vision étendue.

## 6.3 Screenshot of the package “Copula-tomography”

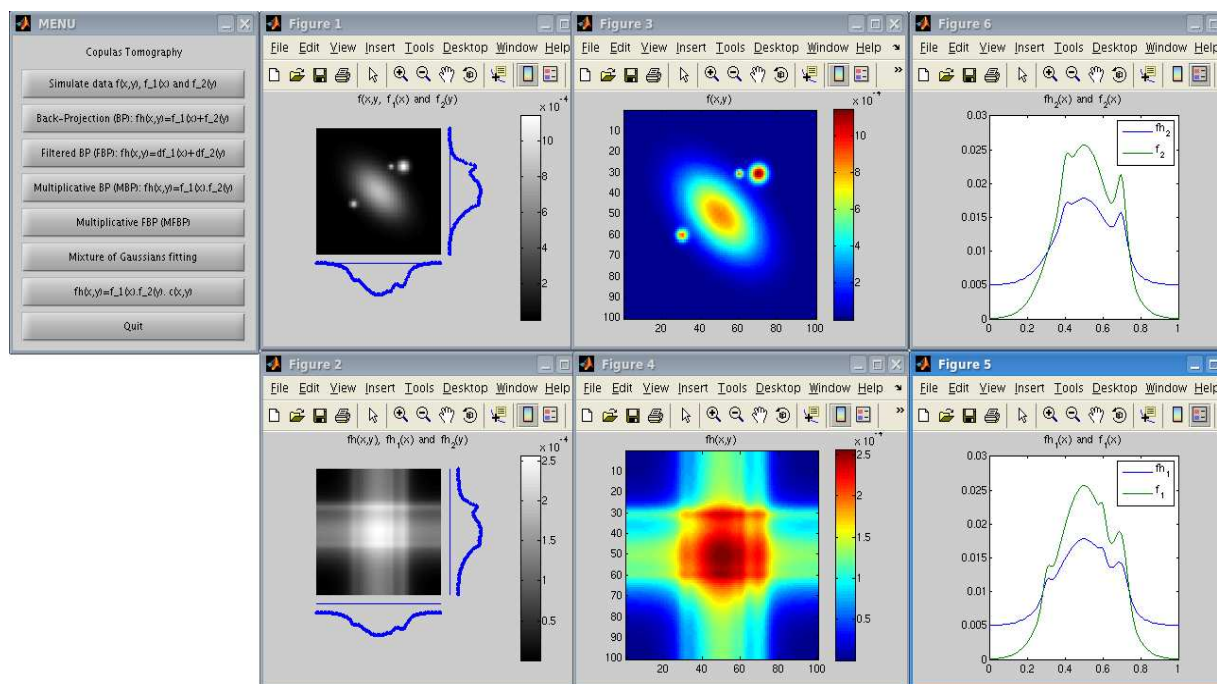


FIGURE 6.5 – Graphical «Menu» and «BP simulation using 4 Gaussians pdf’s : using only 2 projections»

## 6.4 Guide

In this section, we explain how to use this software written to implement the method we have discussed. Version of this code exist in MATLAB, and are freely available by request to the authors. We hope that this package will help the user to enjoy [46] the joy of copula [50] in tomography

To begin using the code, download the file COPULATOMOGRAPHY.zip and extract the directory COPULATOMOGRAPHY. This directory contains the MATLAB ®program.

Below is a short description of the package we have written using Matlab ®which allows users to simulate a tomographic image reconstruction.

### 6.4.1 Menu

The main menu is titled : Copulas Tomography  
It gives a user interface that allows to make selections and choices from the following preset lists :

1. Simulate data :  $f(x,y), f_1(x)$  and  $f_2(y)$

2. Back-Projection (BP) :  $f_h(x, y) = f_1(x) + f_2(y)$
3. Filtered BP (FBP) :  $f_h(x, y) = df_1(x) + df_2(y)$
4. Multiplicative BP (MBP) :  $f_h(x, y) = f_1(x) \cdot f_2(y)$
5. Multiplicative FBP (MFBP) :  $f_h(x, y) = df_1(x) \cdot df_2(y)$
6. Copulas based reconstruction :  $f_h(x, y) = f_1(x) \cdot f_2(y) \cdot c(F_1(x), F_2(y))$
7. Quit

### 6.4.2 Menu List

Let us explain each part of the previous list :

1. **Simulate data  $f(x, y)$ ,  $f_1(x)$  and  $f_2(y)$**

In fact, this list is a pop-up menu which contains a group of nine choices as it follows :

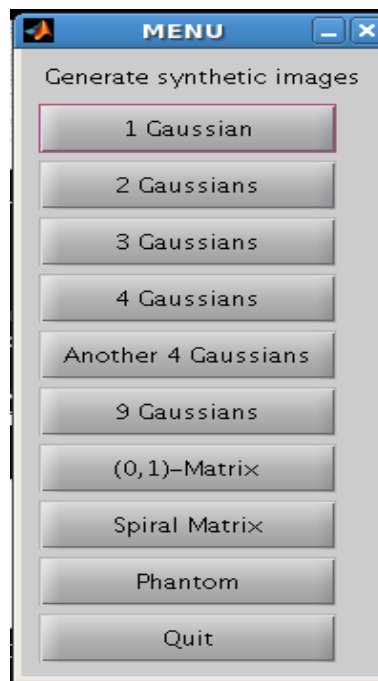


FIGURE 6.6 – MENU :Generate Synthetic images

Those choices are related to the kind of data to be simulated. Of course, one can extend it.

Those data represent  $f(x, y)$ , the original image. We offer many choices to the user. After the selection of the data, clicking on «quit» will close the pop-up menu and allow the user to return to the main menu.

2. **Back-Projection (BP) :  $f_h(x, y) = f_1(x) + f_2(y)$**

This list offers a visualisation of the reconstruction using Back-Projection method.  $f_h(x, y)$  is the image reconstructed.

3. **Filtered BP (FBP) :  $f_h(x, y) = df_1(x) + df_2(y)$**

This list offers a visualisation of the reconstruction using Filtered Back-Projection method.

4. **Multiplicative BP (MBP) :  $f_h(x, y) = f_1(x).f_2(y)$**

This list offers a visualisation of the tomographic image reconstruction using Multiplicative Back-Projection method.

5. **Multiplicative FBP (MFBP)  $f_h(x, y) = df_1(x).df_2(y)$**

This list offers a visualisation of the reconstruction using Multiplicative Filtered Back-Projection method.

6. **Copulas based reconstruction :  $f_h(x, y) = f_1(x).f_2(y).c(F_1(x), F_2(y))$**

Click on this list of the main menu yields to a large number of copula family  $c(F_1(x), F_2(y))$  to reconstruct the original object. We offer the reconstruction through the gaussian copula, the independent copula, and  $\exp(F_1 + F_2)$ . Of course the user can add many other families of copulas.

7. **Quit** : To exit from the main menu.

## 6.5 Conclusion

---

Ce chapitre presente essentiellement quelques résultats des expérimentations faites pour reconstruire une image en tomographie avec deux projections. Nous avons décrit les logiciels écrits en langage Matlab® permettant d'utiliser la méthode que nous avons proposée.

# 7

## Conclusions générales et Perspectives

Nous présentons ici quelques perspectives de recherches basées sur les résultats que nous avons présentés dans ce manuscrit.

### 7.1 La classe de copules entropiques

---

Le chapitre où nous avons présenté les “copules à maximum d’entropie” constitue à notre avis la partie essentielle de cette thèse. Il reste cependant des problèmes ouverts dans plusieurs directions, que nous essayons de présenter dans cette partie perspective. Nous avons montré le lien que la copule pourrait avoir avec différentes formes d’entropie une fois que la forme explicite était mathématiquement possible, par exemple pour la valeur  $q = 2$  de l’index d’entropie de Rényi ou bien l’entropie de Tsallis. Nous avons montré les liens possibles avec la famille de copule de Farlie-Gumbel-Morgenstern (FGM) et les copules archimédiennes. Nous avons vérifié sous quelques conditions et pour certaines valeurs des paramètres des copules qu’on a obtenu, que ces copules sont bien définies dans l’intervalle  $[0, 1]$ .

- Mais quel sera l’intervalle où se situent tous les paramètres pour que la copule soit toujours dans l’intervalle  $[0, 1]$  ?
- Comment peut-on associer chaque classe de copules connues dans la littérature statistique à chacune des différentes entropies connues ?

### 7.2 Tomographie discrète et copule discrète

---

Le rapprochement entre tomographie discrète et copules discrètes donne une autre perspective de recherches approfondies pour faire ressortir tous les liens existants. En partant de l’article publié en 1992 par William F. Darsow, Bao Gia Nguyen et Elwood T. Olsen qui ont fait la connection entre les copules et les processus Markoviens [79]. Par la suite certains travaux de recherches s’en sont inspirés. Le lien avec l’algèbre des matrices sont également des perspectives ouvertes de travaux dans le cadre de la tomographie, et plus précisément dans la résolution des problèmes de la tomographie discrète.



### 7.3 Extension des copules : notion de $q$ -copule

---

Les copules sont construites à partir des lois de distributions de probabilités classiques. Une perspective serait de voir s'il n'y aurait pas un chemin vers l'extension de la notion de copule dans le cadre des "statistiques de Tsallis" [64].

Par exemple si on considère  $F_q^i$  des fonctions de répartition  $q$ -Gaussiennes, peut-on avoir aussi  $C_q$  qui pourraient être vue comme une "copule  $q$ -Gaussienne". On sait que le cas particulier de la valeur de  $q = 1$  (voir par exemple [90]) donne la fonction de répartition gaussienne classique, donc  $C_1$  est la copule classique gaussienne.

Mais qu'en sera-t-il de  $q \neq 1$  ?

Existe-t-il un "théorème" qui serait une extension du théorème de Sklar [5], et qui permettra d'écrire pour des lois conjointes  $F_q(x_1, x_2)$  dont les fonctions marginales données  $F_q^i$  pour  $i = 1, 2$  sont définies au sens de la "statistique de Tsallis" (voir par exemple [91]).

Ce "théorème" permettrait d'écrire

$$F_q(x_1, x_2) \equiv C_q(F_q^1(x_1), F_q^2(x_2)). \quad (7.1)$$

### 7.4 Copules maximales de divergences et d'entropies

---

Dans notre rapport, le cas particulier pour l'indice d'entropie  $q = 2$ , s'est avéré important et nous a permis de montrer l'existence de nouvelles familles de copules et les liens qu'il y a avec d'autres copules déjà connues.

Nous savons qu'il existe plusieurs types d'entropie et de mesures de divergences ayant des applications fondamentales [62]. Une autre perspective de travaux serait de chercher à caractériser les copules maximales de ces autres entropies (Par exemple, l'entropie de Burg) ou de divergences sous les contraintes (4.14) et qui peuvent fournir des solutions explicites.

### 7.5 Autres points

---

- Extension de la méthode que nous avons proposée dans le chapitre 4 pour des lois sur  $\mathbb{R}$  et  $\mathbb{R}^+$  au lieu de l'intervalle  $[0, 1]$ .
- Extension si on rajoute d'autres contraintes en plus des marginales, par exemple les moments ou le support.
- Extension extension au cas où on a plus de deux projections.
- Interprétation statistique d'une projection à l'angle  $\pi/4$ .

### 7.6 Conclusions générales

---

Dans ce travail présenté, nous avons :

- Établi un lien entre la notion de copule et la tomographie.
- Décrit un problème commun pour l'inversion
- Proposé une méthode d'inversion en utilisant les entropies de
  - Shannon
  - Rényi
  - Tsallis
- Découvert de nouvelles copules en utilisant l'entropie de Tsallis avec  $q = 2$  et les densités marginales Bêta et Kumaraswamy.
- Développé un outil de simulation en Matlab.



# A

## Sur la preuve des densités de probabilité maximisant des entropies

### A.1 Problème d'optimisation continue

---

Nous avons cherché une approche rigoureuse qui pourrait permettre à étendre l'équation (A.55) pour qu'elle soit identifiable à la méthode classique des multiplicateur de Lagrange, ce qui pourrait alors nous permettre d'enlever sans ambiguïté mathématique les deux parenthèses de l'expression "méthode de multiplicateur", car nous savons bien que la méthode de multiplicateur de Lagrange ne peut s'appliquer que dans le cas où les multiplicateurs de Lagrange sont en nombre fini, ce qui n'est pas le cas des fonctions  $\lambda_i$  dans l'équation (A.55). Une piste de réflexion qui pourrait permettre de contourner ce que nous avons trouvé, mais qui est seulement fondée sur l'entropie de Shannon se trouve dans [68]. Nous exposons brièvement un cadre général, dans un espace de fonction  $L_1$ .

Comme nous l'avons déjà noté, il y a en effet, une infinité de contraintes dans l'expression des équations (4.14), nous avons intuitivement par une méthode heuristique de recherche de la forme de la densité conjointe optimale (sous certaines hypothèses) aboutit à la découverte de nos familles de copules. Le terme heuristique est utilisé car nous savons bien que la méthode classique de Lagrange n'était pas utilisée. Mais une analyse de la méthode qui est définie dans [68], qui est aussi utilisée dans [69] sous le terme de "problème d'optimisation continu" est la mieux adaptée pour le problème (A.55), pourrait guider par les grandes étapes suivantes que nous décrirons ci-après et conduire à la solution plus rigoureuse permettant de conclure à notre résultat. Si nous considérons les équations (4.14) dans une classe un peu plus générale comme celles des fonctions mesurables. Nous rappelons brièvement qu'une fonction définie sur un espace mesuré  $\Omega$  et à valeurs dans  $\mathbb{R}$  ou  $\mathcal{C}$  est dite de "carré sommable" ou de "carré intégrable" si elle appartient à l'espace  $L^2(\Omega)$ .

Par exemple, une fonction mesurable  $f$  de  $\mathbb{R}$  dans  $\mathcal{C}$  est de carré sommable lorsque l'intégrale suivante (au sens de l'intégrale de Lebesgue)

$$S = \int_{-\infty}^{+\infty} |f(x)|^2 dx$$

converge, c'est-à-dire si elle existe et correspond à un nombre fini.

On peut alors remplacer les expressions pour tout  $x_1$  et pour tout  $x_2$  dans les équations

## ANNEXE A. SUR LA PREUVE DES DENSITÉS DE PROBABILITÉ MAXIMISANT DES ENTROPIES

---

tions (4.14) par “presque partout” (p.p), et nous pourrions dans ce nouvel espace de fonctions conclure au problème dit “d’optimisation continue”[68, 69]

L’idée sera alors de définir une application linéaire  $A$ , pour qu’alors notre problème (4.14) soit traité dans cet espace de fonctions.

**Définition A.1.0.1.** *On peut définir, dans le cas où l’espace considéré est  $L_1$  en suivant le schéma dans [68]*

$$A : L_1(I \times I) \longrightarrow L_1(I) \times L_1(I) \times L_1(I)$$

$$f \longmapsto Af = \left( \int_0^1 U(x_1)dx_1, \int_0^1 U(x_2)dx_2, \int_0^1 \int_0^1 f(x_1, x_2)dx_1 dx_2 \right)$$

Tout le problème serait de bien choisir les fonctions  $U(x_1)$  et  $U(x_2)$  (qui seront normalisées d’une certaine manière, pour que les contraintes dans les équations (4.14) soient finies c’est-à-dire l’application  $A(f) = (1, 1, 1)$ ).

Après l’étape précédente, on pourrait alors déduire le théorème suivant (toujours en suivant la démonstration dans [68] ) :

**Théorème A.1.1.** *Il existe une solution unique au problème (4.14) qui est de la forme*

$$f(x_1, x_2) = f_2(x_2) + f_1(x_1) - 1$$

avec  $f$  continue.

**Preuve 2.** *La proposition 1 [page 10, [69]] peut être adaptée pour prouver l’existence de  $f$  et l’argumentation pour prouver le Théorème 6 [69] achève la partie concernant la continuité de  $f$ .*

*En remarquant simplement que pour notre choix de l’indice d’entropie  $q = 2$ , cette valeur est supérieure à un et l’entropie de Tsallis (ou de Rényi) est concave. Et donc cette concavité entraîne que l’expression “minimisation” dans la proposition 1 [page 10, [69]] est remplacée par l’expression “maximisation d’entropie”. En effet, le problème d’optimisation que nous avons considéré et qui a conduit à notre nouvelle famille de copule et qui est appelé “problème d’optimisation mesurable” est le cas continu des problèmes d’optimisation d’entropie de Shannon [page 273, [68]], où l’auteur considère deux espaces de mesure finie  $(X_1, dx_1)$  et  $(X_2, dx_2)$  avec  $0 \leq f_1 \in L_1(X_1, dx_1)$  et  $0 \leq f_2 \in L_2(X_2, dx_2)$ , nous pouvons conclure la preuve du théorème que nous avons énoncé en adaptant la procédure, toujours dans [68] à la page 273-274, si nous remarquons que  $k \equiv 1$ ,  $\phi(f) \equiv J_i(f)$ , et  $A(f) \equiv \mathcal{L}_i(f)$ .*

## A.2 Autre preuve à partir de la divergence $(a, \lambda)$

---

Dans cette section, nous montrons que les formules dérivées dans les équations (4.27) et (4.29) et provenant de la procédure ayant une ressemblance avec la méthode des multiplicateurs de Lagrange, et dans le cas d’intervalle fini peuvent être étendues à un intervalle infini et conduire à l’unique solution maximale  $q$ -entropie.

En effet, nous savons et nous avons aussi réalisé que la technique du multiplicateur de Lagrange nécessitent des conditions (qualification de la contrainte, conditions de Karush-Kuhn-Tucker (KKT)), et que l'extension à un cas infini, bien que raisonnable, n'est pas évidente. Par conséquent, nous proposons ici une approche différente, en commençant par l'utilisation d'une technique similaire avec la méthode de Lagrange qui nous présente une solution potentielle. Ensuite, nous prouvons que cette solution potentielle est effectivement la meilleure.

Le problème est comme suit. Soit une fonctionnelle d'entropie  $J_i(f)$  définie pour le cas bivarié ou d'une densité conjointe  $f$  suivant la mesure de Lebesgue, trouver cette densité à partir de ses densités marginales qui maximisent l'entropie fonctionnelle :

$$\begin{cases} \max & J_i(f) \\ \text{s.t.} & f_1(x_1) = \int f(x_1, x_2) dx_2 \\ & f_2(x_2) = \int f(x_1, x_2) dx_1 \\ & f \geq 0, \iint f(x_1, x_2) dx_1 dx_2 = 1 \end{cases}$$

Dès lors que  $J_i(f)$  est concave et les contraintes sont linéaires, la solution est unique. Nous allons considérer les fonctionnelles qui sont les entropies de Rényi, Tsallis et Shannon. Les deux premières entropies étant liés par une fonction simple monotone, la maximisation sous contrainte de celles-ci donnera la même solution. Enfin, la solution pour la solution pour l'entropie de Shannon est obtenue comme cas particulier ( $q = 1$ ).

Les définitions des divergences  $(a, \lambda)$  ont été introduites en [92]. Elles permettent de retrouver des cas particuliers d'extension des divergences  $(\alpha, \beta)$  suggérés dans [93] Soient deux paramètres  $a$  et  $\lambda > 0$  et deux densités  $f$  et  $g$ , la divergence  $(a, \lambda)$ -Rényi est

$$D_{a,\lambda}^{(R)}(f||g) = \frac{1}{a - \lambda} \log \frac{\left[ \int f(x)^a g(x)^{\lambda-a} d\mu(x) \right]^\lambda}{\left[ \int f(x)^\lambda d\mu(x) \right]^a \left[ \int g(x)^\lambda d\mu(x) \right]^{\lambda-a}}, \quad (\text{A.1})$$

à condition que les différentes intégrales existent.

De la même manière, la divergence  $(a, \lambda)$ -Tsallis est la quantité

$$D_{a,\lambda}^{(T)}(f||g) = \frac{1}{a - \lambda} \left( \frac{\left[ \int f(x)^a g(x)^{\lambda-a} d\mu(x) \right]^\lambda}{\left[ \int f(x)^\lambda d\mu(x) \right]^a \left[ \int g(x)^\lambda d\mu(x) \right]^{\lambda-a}} - 1 \right) \quad (\text{A.2})$$

à condition que les différentes intégrales existent. Ces deux divergences sont facilement liées aux divergences classiques, en utilisant le concept des "escort distributions"[94] , alors, nous avons le résultat suivant (voir [92]) :

**Proposition A.2.1.** *Si  $f$  et  $g$  sont deux densités telles que les intégrales soient finies, avec  $a$  et  $\lambda > 0$ , alors  $D_{a,\lambda}^{(\cdot)}(f||g) \geq 0$ , avec égalité si et seulement si  $f = g$ .*

## ANNEXE A. SUR LA PREUVE DES DENSITÉS DE PROBABILITÉ MAXIMISANT DES ENTROPIES

---

Nous pouvons utiliser ici  $a = 1$ , et nous prenons  $\lambda = q$ , qui donnent la divergence de Tsallis et aussi le cas  $a = q$  et  $\lambda = 1$  la divergence de Rényi.

Maintenant basé sur le fait de l'équation (4.15) et si nous réécrivons le résultat (4.27) que nous avons montré pour l'entropie de Rényi

$$\begin{aligned} f^{q-1}(x_1, x_2) &= \left( \iint_{\mathbf{I}^2} f^q(x_1, x_2) dx_1 dx_2 \right) \frac{1-q}{q} (\lambda_1(x_1) + \lambda_2(x_2) + \lambda_0) \\ f(x_1, x_2) &= \left( \left( \iint_{\mathbf{I}^2} f^q(x_1, x_2) dx_1 dx_2 \right) \frac{1-q}{q} \right)^{\frac{1}{q-1}} (\lambda_1(x_1) + \lambda_2(x_2) + \lambda_0)^{\frac{1}{q-1}}. \end{aligned} \quad (\text{A.3})$$

**Remarque 8.** Nous rappelons ici que, lorsque nous avons montré le résultat dans le chapitre 4, nous supposons que  $\iint_{\mathbf{I}^2} f^q(x_1, x_2) dx_1 dx_2 = 1$ .

De même le résultat pour l'entropie de Tsallis (4.29)

$$f(x_1, x_2) = \frac{1-q}{q} (\lambda_1(x_1) + \lambda_2(x_2) + \lambda_0)^{\frac{1}{q-1}}. \quad (\text{A.4})$$

Nous pouvons introduire l'expression de la solution de  $q$ -entropie maximale pour le cas infini, c'est-à-dire  $(x_1, x_2) \in \mathbf{R}^2$ .

Notons

$$g_{ME}(x_1, x_2) = (\lambda_0 + \lambda_1(x_1) + \lambda_2(x_2))^{\frac{1}{q-1}} \quad (\text{A.5})$$

la solution  $q$ -entropique supposée, où  $\lambda_0, \lambda_1(x_1), \lambda_2(x_2)$  sont choisies telle que  $g_{ME}$  est normalisé à un et les densités marginales données

**Proposition A.2.2.** Pour toute densité  $f(x_1, x_2)$  avec les deux marginales données  $f_1(x_1)$  et  $f_2(x_2)$ , si  $g_{ME}$  est défini comme ci-dessus dans (A.5), alors  $H_q[g_{ME}] \geq H_q[f]$ , avec  $H_q$  désignant les entropies de Rényi ou de Tsallis.

*Démonstration.* Évaluons d'abord  $\iint f(x_1, x_2)g_{ME}(x_1, x_2)^{q-1}dx_1dx_2$ . Nous avons

$$\begin{aligned} \iint f(x_1, x_2)g_{ME}(x_1, x_2)^{q-1}dx_1dx_2 &= \iint f(x_1, x_2) (\lambda_0 + \lambda_1(x_1) + \lambda_2(x_2)) dx_1dx_2 \\ &= \lambda_0 + \iint f(x_1, x_2)\lambda_1(x_1)dx_1dx_2 + \iint f(x_1, x_2)\lambda_2(x_2)dx_1dx_2, \\ &= \lambda_0 + \int f_1(x_1)\lambda_1(x_1)dx_1 + \int f_2(x_2)\lambda_2(x_2)dx_2, \end{aligned}$$

où nous avons utilisé le fait que  $f(x_1, x_2)$  a les densités marginales  $f_1(x_1)$  et  $f_2(x_2)$ . Évidemment, si  $g_{ME}$  a les mêmes marginales, nous obtenons

$$\begin{aligned} \iint g_{ME}(x_1, x_2)g_{ME}(x_1, x_2)^{q-1}dx_1dx_2 &= \iint g_{ME}(x_1, x_2)^q dx_1dx_2 \\ &= \lambda_0 + \int f_1(x_1)\lambda_1(x_1)dx_1 + \int f_2(x_2)\lambda_2(x_2)dx_2, \end{aligned}$$

avec  $f = g_{ME}$ . Par conséquent, nous avons l'identité

$$\iint f(x_1, x_2) g_{ME}(x_1, x_2)^{q-1} dx_1 dx_2 = \iint g_{ME}(x_1, x_2)^q dx_1 dx_2. \quad (\text{A.6})$$

Par la positivité des  $(a, \lambda)$ -divergences, nous obtenons alors

$$D_{1,q}^{(T)}(f||g_{ME}) = \frac{1}{1-q} \left( \frac{\left[ \int f(x) g_{ME}(x)^{q-1} d\mu(x) \right]^q}{\left[ \int f(x)^q d\mu(x) \right] \left[ \int g(x)^q d\mu(x) \right]^{q-1}} - 1 \right) \geq 0;$$

puis en utilisant l'égalité dans (A.6) et en simplifiant, il vient

$$D_{1,q}^{(T)}(f||g_{ME}) = \frac{1}{1-q} \left( \frac{\left[ \int g_{ME}(x)^q d\mu(x) \right]}{\left[ \int f(x)^q d\mu(x) \right]} - 1 \right) \geq 0,$$

qui, à son tour implique que  $H_q[g_{ME}] \geq H_q[f]$ . Le signe d'égalité est obtenu en cas d'égalité dans (A.2), c'est-à-dire si et seulement si  $f = g_{ME}$ .  $\square$

### A.2.1 Remarque sur le cas indépendant

En général, le cas indépendant

$$\lim_{q \rightarrow 1} g_{ME}(x_1, x_2) = f_1(x_1) f_2(x_2), \quad (\text{A.7})$$

est toujours vérifié car nous savons que nous avons affaire à l'entropie de Shannon car mentionné ci-dessus dans le cas  $q = 1$ .

Pour conclure cette partie, il y a quelques remarques si nous considérons d'abord  $q \neq 1$  et que nous cherchons les expressions explicites de  $A_0(q)$  et  $\tilde{\lambda}_i$ .

Nous avons

$$g_{ME}(x_1, x_2) = (\lambda_0 + \lambda_1(x_1) + \lambda_2(x_2))^{\frac{1}{q-1}}, \quad (\text{A.8})$$

$$g_{ME}(x_1, x_2) = (\lambda_0)^{\frac{1}{q-1}} \left( 1 + \frac{\lambda_1(x_1)}{\lambda_0} + \frac{\lambda_2(x_2)}{\lambda_0} \right)^{\frac{1}{q-1}}. \quad (\text{A.9})$$

En supposons que l'expression explicite où converge  $g_{ME}$  peut être obtenue alors les étapes devraient être

$$g_{ME}(x_1, x_2) = A_0(q) \left( 1 + (q-1)\tilde{\lambda}_1(x_1) + (q-1)\tilde{\lambda}_2(x_2) \right)^{\frac{1}{q-1}}, \quad (\text{A.10})$$

qui peut être écrite

$$g_{ME}(x_1, x_2) = A_0(q) \exp \left( \frac{1}{q-1} \log \left( 1 + (q-1)\tilde{\lambda}_1(x_1) + (q-1)\tilde{\lambda}_2(x_2) \right) \right). \quad (\text{A.11})$$



## ANNEXE A. SUR LA PREUVE DES DENSITÉS DE PROBABILITÉ MAXIMISANT DES ENTROPIES

---

En posant  $x = (1 - q) \left( \tilde{\lambda}_1(x_1) + \tilde{\lambda}_2(x_2) \right)$ .

Nous savons que  $x \rightarrow 0$ ,  $\log(1 + x) \sim x$  pour  $x > -1$ , alors quand  $q \rightarrow 1$  nous avons  $\log \left( 1 + (q - 1) \left( \tilde{\lambda}_1(x_1) + \tilde{\lambda}_2(x_2) \right) \right) \sim (q - 1) \left( \tilde{\lambda}_1(x_1) + \tilde{\lambda}_2(x_2) \right)$

À partir de l'équation (A.11)

$$\lim_{q \rightarrow 1} g_{ME}(x_1, x_2) = A_0(q) \exp \left( \tilde{\lambda}_1(x_1) + \tilde{\lambda}_2(x_2) \right) \quad (\text{A.12})$$

$$\lim_{q \rightarrow 1} g_{ME}(x_1, x_2) = A_0(q) \exp \left( \tilde{\lambda}_1(x_1) \right) \exp \left( \tilde{\lambda}_2(x_2) \right), \quad (\text{A.13})$$

Pour continuer cette analyse, nous voyons ici qu'il faut explicitement trouver l'expression de  $g_{ME}$  en fonction uniquement des densités marginales  $f_i(x_i)$ .

Si  $g_{ME}$  converge en loi vers la distribution gaussienne généralisée, la  $q$ -Gaussienne, la distribution  $q$ -Student ou bien l'une des  $q$ -distributions parmi celles qui sont citées en [91] alors il est possible d'obtenir les densités marginales  $f_i(x_i)$  et donc d'avoir aussi  $A_0(q)$  de telle sorte que la limite quand  $q \rightarrow 1$ , donne  $g_{ME} = \prod_i f_i(x_i)$ .

Dans le cas particulier  $1 < q < 2$ , il y a la convergence en loi vers la  $q$ -Gaussienne et donc possibilité d'obtenir explicitement la constante  $A_0(q)$  qui permet de conclure au résultat dit de  $q$ -indépendance qui se base sur le  $q$ -Théorème Central Limite [voir [95]].

Mais nous remarquons aussi que la constante de normalisation peut être obtenue à l'aide de l'équation (A.9) en posant  $\lambda_i(x_i) = x_i^2$  (c'est-à-dire  $g_{ME}$  est une  $q$ -gaussienne bivariée) et on peut extraire  $A_0(q)$  via le résultat de [96] en faisant le changement adéquat dans l'expression  $\mathcal{C}_N(q)$ .

### A.3 Certains détails de calculs

---

### A.4 Notation

---

- $\ln(x)$  : la fonction logarithme de  $x$  ;
- $\exp(x)$  : la fonction exponentielle de  $x$  ;
- $L$  : un nombre réel (ou  $\pm\infty$ ) ;
- $c$  : un nombre réel ;
- $N(x)$  : la fonction au numérateur ;
- $D(x)$  : la fonction au dénominateur ;
- $N'(x) = \frac{dN(x)}{dx}$  : dérivée de la fonction  $N$  par rapport à  $x$  ;
- $D'(x) = \frac{dD(x)}{dx}$  : dérivée de la fonction  $D$  par rapport à  $x$  ;
- $S$  : entropie de Shannon ;
- $q$  : indice d'entropie
- $T_q$  : entropie de Tsallis d'indice d'entropie  $q$  ;
- $R_q$  : entropie de Rényi d'indice d'entropie  $q$  ;

- $f(x_1, x_2)$  : densité conjointe de probabilité
- $\psi(q) = \iint f^q(x_1, x_2) dx_1 dx_2$ .

### A.4.1 Question

Dans la démonstration de la preuve au chapitre 4, détaille de la manière la plus claire possible le passage de  $f(x_1, x_2) = -\frac{1}{2}(\lambda_1(x_1) + \lambda_2(x_2) + \lambda_0)$  à  $f(x_1, x_2) = f_1(x_1) + f_2(x_2) - 1$ .

### A.4.2 Réponse 1

$$f(x_1, x_2) = -\frac{1}{2}(\lambda_1(x_1) + \lambda_2(x_2) + \lambda_0). \quad (\text{A.14})$$

L'équation (A.14) doit satisfaire les contraintes imposées suivantes

$$\left\{ \begin{array}{ll} \int_0^1 f(x_1, x_2) dx_2 = f_1(x_1), & \forall x_1 \quad (i) \\ \int_0^1 f(x_1, x_2) dx_1 = f_2(x_2), & \forall x_2 \quad (ii) \\ \int_0^1 \int_0^1 f(x_1, x_2) dx_1 dx_2 = 1, & (iii) \\ f(x_1, x_2) \geq 0, & (iv) \end{array} \right. \quad (\text{A.15})$$

Le cas (iii) dans la condition (A.15) conduit à :

$$\begin{aligned} \int_0^1 \int_0^1 f(x_1, x_2) dx_1 dx_2 &= \int_0^1 \int_0^1 \left[ -\frac{1}{2}(\lambda_1(x_1) + \lambda_2(x_2) + \lambda_0) \right] dx_1 dx_2 \\ &= -\frac{1}{2} \left( \int_0^1 \int_0^1 \lambda_1(x_1) dx_1 dx_2 + \int_0^1 \int_0^1 \lambda_2(x_2) dx_1 dx_2 + \int_0^1 \int_0^1 \lambda_0 dx_1 dx_2 \right) \\ &= -\frac{1}{2} \left( \int_0^1 \lambda_1(x_1) dx_1 \int_0^1 dx_2 + \int_0^1 \lambda_2(x_2) dx_2 \int_0^1 dx_1 + \lambda_0 \int_0^1 \int_0^1 dx_1 dx_2 \right) \\ &= -\frac{1}{2} \left( \int_0^1 \lambda_1(x_1) dx_1 + \int_0^1 \lambda_2(x_2) dx_2 + \lambda_0 \right) = 1 \quad (\text{A.16}) \end{aligned}$$

On déduit de l'équation (A.16) l'expression de  $\lambda_0$  :

$$\lambda_0 = -2 - \int_0^1 \lambda_1(x_1) dx_1 - \int_0^1 \lambda_2(x_2) dx_2. \quad (\text{A.17})$$

## ANNEXE A. SUR LA PREUVE DES DENSITÉS DE PROBABILITÉ MAXIMISANT DES ENTROPIES

---

Le cas (i) dans la condition (A.15) conduit à :

$$\int_0^1 f(x_1, x_2) dx_2 = \int_0^1 \left[ -\frac{1}{2}(\lambda_1(x_1) + \lambda_2(x_2) + \lambda_0) \right] dx_2 \quad (\text{A.18})$$

$$= -\frac{1}{2} \left[ \int_0^1 \lambda_1(x_1) dx_2 + \int_0^1 \lambda_2(x_2) dx_2 + \int_0^1 \lambda_0 dx_2 \right] \quad (\text{A.19})$$

$$= -\frac{1}{2} \left( \lambda_1(x_1) \int_0^1 dx_2 + \int_0^1 \lambda_2(x_2) dx_2 + \lambda_0 \int_0^1 dx_2 \right) \quad (\text{A.20})$$

$$= -\frac{1}{2} \left( \lambda_1(x_1) + \int_0^1 \lambda_2(x_2) dx_2 + \lambda_0 \right) = f_1(x_1) \quad (\text{A.21})$$

On déduit de l'équation (A.21) l'expression de  $\lambda_1(x_1)$  :

$$\lambda_1(x_1) = -2f_1(x_1) - \int_0^1 \lambda_2(x_2) dx_2 - \lambda_0. \quad (\text{A.22})$$

Avec l'expression de  $\lambda_0$  obtenu en (A.17), l'équation (A.22) devient

$$\begin{aligned} \lambda_1(x_1) &= -2f_1(x_1) - \int_0^1 \lambda_2(x_2) dx_2 - \left( -2 - \int_0^1 \lambda_1(x_1) dx_1 - \int_0^1 \lambda_2(x_2) dx_2 \right) \\ &= -2f_1(x_1) + \int_0^1 \lambda_1(x_1) dx_1 + 2. \end{aligned} \quad (\text{A.23})$$

De même en utilisant le cas (ii) dans la condition (A.15) conduit à :

$$\lambda_2(x_2) = -2f_2(x_2) + \int_0^1 \lambda_2(x_2) dx_2 + 2. \quad (\text{A.24})$$

En substituant dans l'équation (A.14) les expressions de  $\lambda_0$ ,  $\lambda_1(x_1)$  et  $\lambda_2(x_2)$  obtenues respectivement dans les équations (A.17), (A.23) et (A.24) , on a alors le résultat final :

$$f(x_1, x_2) = f_1(x_1) + f_2(x_2) - 1. \quad (\text{A.25})$$

### A.5 Question 2

---

Quand  $q \rightarrow 1$ , **montrer de manière détaillée** que l'entropie de Rényi :

$$R_q(f) = \frac{1}{1-q} \ln \left( \iint f^q(x_1, x_2) dx_1 dx_2 \right), q \geq 0, q \neq 1 \quad (\text{A.26})$$

et l'entropie de Tsallis :

$$T_q(f) = \frac{1}{q-1} \left( 1 - \iint f^q(x_1, x_2) dx_1 dx_2 \right), q \geq 0, q \neq 1 \quad (\text{A.27})$$

**coïncident bien** à l'entropie de Shannon :

$$S(f) = - \iint f(x_1, x_2) \ln f(x_1, x_2) dx_1 dx_2. \quad (\text{A.28})$$

## A.6 Réponse 2

Pour répondre à la question 1, nous allons utiliser la règle de l'Hôpital<sup>1</sup> :

### A.6.1 Règle de l'Hôpital

Si  $\lim_{x \rightarrow L} N(x) = \lim_{x \rightarrow L} D(x) = 0$  (ou  $\pm\infty$ ) et  $\lim_{x \rightarrow L} \frac{N'(x)}{D'(x)} = \mathbf{c}$  existe, alors

$$\lim_{x \rightarrow L} \frac{N(x)}{D(x)} = \lim_{x \rightarrow L} \frac{N'(x)}{D'(x)} = \mathbf{c}. \quad (\text{A.29})$$

## A.7 Quelques expressions utiles

Dans tout ce qui suit la fonction  $f(x_1, x_2)$  est **fixée** et la variable est  $q$ .

$$f^q(x_1, x_2) = \exp[\ln(f^q(x_1, x_2))] = \exp[q \ln(f(x_1, x_2))], \quad (\text{A.30})$$

on en déduit l'expression de la dérivée par rapport à  $q$  :

$$\frac{d[f^q(x_1, x_2)]}{dq} = \ln(f(x_1, x_2)) \exp(q \ln(f(x_1, x_2))). \quad (\text{A.31})$$

On rappelle aussi que pour n'importe quelle fonction  $\psi(q)$  dérivable, la dérivée de la fonction logarithme de  $\psi(q)$  par rapport à  $q$  :

$$\frac{d[\ln(\psi(q))]}{dq} = \frac{\psi'(q)}{\psi(q)}. \quad (\text{A.32})$$

### A.7.1 Expression des dérivées

$$\begin{aligned} [\psi(q)]' &= \frac{d}{dq} \left[ \iint f^q(x_1, x_2) dx_1 dx_2 \right] \\ &= \iint \frac{d[f^q(x_1, x_2)]}{dq} dx_1 dx_2 \end{aligned} \quad (\text{A.33})$$

et en utilisant (A.31), il vient que

$$[\psi(q)]' = \iint \ln(f(x_1, x_2)) \exp(q \ln(f(x_1, x_2))) dx_1 dx_2. \quad (\text{A.34})$$

1. Weisstein, Eric W. "L'Hospital's Rule." From MathWorld—A Wolfram Web Resource. <http://mathworld.wolfram.com/LHospitalsRule.html>

### A.7.2 Montrons que $R_q(f) = S(f)$ quand $q \rightarrow 1$

On pose d'après l'équation (A.26) qui donne l'entropie de Rényi :

$$N(q) = \ln \left( \iint f^q(x_1, x_2) dx_1 dx_2 \right) = \ln[\psi(q)], \quad (\text{A.35})$$

et

$$D(q) = 1 - q. \quad (\text{A.36})$$

et d'après l'équation (A.62), il vient

$$N'(q) = \frac{[\psi(q)]'}{\psi(q)}. \quad (\text{A.37})$$

Il vient d'après les équations (A.59) et (A.63) que :

$$N'(q) = \frac{\iint \ln(f(x_1, x_2)) \exp(q \ln(f(x_1, x_2))) dx_1 dx_2}{\iint \exp(q \ln(f(x_1, x_2))) dx_1 dx_2}. \quad (\text{A.38})$$

En utilisant (A.36) on a aussi :

$$D'(q) = -1. \quad (\text{A.39})$$

### A.7.3 Calcul des limites

En utilisant les équations (A.29), (A.63), (A.64) on a :

$$\begin{aligned} \lim_{q \rightarrow 1} \frac{N(q)}{D(q)} &= \lim_{q \rightarrow 1} \frac{N'(q)}{D'(q)} \\ &= \lim_{q \rightarrow 1} \frac{\iint \ln(f(x_1, x_2)) [\exp(q \ln(f(x_1, x_2)))] dx_1 dx_2}{\iint \exp(q \ln(f(x_1, x_2))) dx_1 dx_2} \\ &= \lim_{q \rightarrow 1} \frac{\iint \ln(f(x_1, x_2)) [\exp(\ln(f(x_1, x_2)))] dx_1 dx_2}{-1} \\ &= - \frac{\iint \ln(f(x_1, x_2)) [\exp(\ln(f(x_1, x_2)))] dx_1 dx_2}{\iint \exp(\ln(f(x_1, x_2))) dx_1 dx_2} \\ &= - \frac{\iint \ln(f(x_1, x_2)) [\exp(\ln(f(x_1, x_2)))] dx_1 dx_2}{\iint \exp(\ln(f(x_1, x_2))) dx_1 dx_2}. \end{aligned} \quad (\text{A.40})$$

En remarquant que  $\exp(\ln(f(x_1, x_2))) = f(x_1, x_2)$ , on a d'après l'équation (A.40) :

$$\lim_{q \rightarrow 1} \frac{N(q)}{D(q)} = - \frac{\iint \ln(f(x_1, x_2)) [f(x_1, x_2)] dx_1 dx_2}{\iint f(x_1, x_2) dx_1 dx_2} \quad (\text{A.41})$$

et aussi que  $f(x_1, x_2)$  est une densité conjointe de probabilité c'est-à-dire

$$\iint f(x_1, x_2) dx_1 dx_2 = 1,$$

on a finalement d'après l'équation (A.41)

$$\lim_{q \rightarrow 1} \frac{N(q)}{D(q)} = - \iint \ln(f(x_1, x_2)) [f(x_1, x_2)] dx_1 dx_2 \quad (\text{A.42})$$

$$= - \iint f(x_1, x_2) \ln(f(x_1, x_2)) dx_1 dx_2 \quad (\text{A.43})$$

$$= S(f) \quad (\text{A.44})$$

#### A.7.4 Montrons que $T_q(f) = S(f)$ quand $q \rightarrow 1$

On pose d'après l'équation (A.27) qui donne l'entropie de Tsallis :

$$N(q) = 1 - \iint f^q(x_1, x_2) dx_1 dx_2 = 1 - \psi(q), \quad (\text{A.45})$$

et

$$D(q) = q - 1 \Rightarrow D'(q) = 1. \quad (\text{A.46})$$

D'après l'équation (A.45), il vient

$$N'(q) = - [\psi(q)]' \quad (\text{A.47})$$

Les équations (A.59) et (A.63) donnent :

$$N'(q) = - \iint \ln(f(x_1, x_2)) \exp(q \ln(f(x_1, x_2))) dx_1 dx_2 \quad (\text{A.48})$$

On obtient alors

$$\lim_{q \rightarrow 1} \frac{N(q)}{D(q)} = \lim_{q \rightarrow 1} \frac{N'(q)}{D'(q)} \quad (\text{A.49})$$

$$= \lim_{q \rightarrow 1} \left\{ - \iint \ln(f(x_1, x_2)) [\exp(q \ln(f(x_1, x_2)))] dx_1 dx_2 \right\} \quad (\text{A.50})$$

$$= - \iint \lim_{q \rightarrow 1} \{ \ln(f(x_1, x_2)) [\exp(q \ln(f(x_1, x_2)))] \} dx_1 dx_2 \quad (\text{A.51})$$

$$= - \iint \ln(f(x_1, x_2)) [\exp(\ln(f(x_1, x_2)))] dx_1 dx_2. \quad (\text{A.52})$$

## ANNEXE A. SUR LA PREUVE DES DENSITÉS DE PROBABILITÉ MAXIMISANT DES ENTROPIES

---

En remarquant que  $\exp(\ln(f(x_1, x_2))) = f(x_1, x_2)$

$$\lim_{q \rightarrow 1} \frac{N(q)}{D(q)} = - \iint f(x_1, x_2) \ln(f(x_1, x_2)) dx_1 dx_2 \quad (\text{A.53})$$

$$= S(f) \quad (\text{A.54})$$

### A.7.5 Question 3

$$\begin{aligned} \mathcal{L}_i(f) = & J_i(f) + \lambda_0 \left( 1 - \iint f(x_1, x_2) dx_1 dx_2 \right) + \int \lambda_1(x_1) \left( f_1(x_1) - \int f(x_1, x_2) dx_2 \right) dx_1 \\ & + \int \lambda_2(x_2) \left( f_2(x_2) - \int f(x_1, x_2) dx_1 \right) dx_2, \quad \text{pour } i=1, 2, 3, 4 \end{aligned} \quad (\text{A.55})$$

En considerant dans l'équation (A.55) le cas  $i = 4$  et donc

$$J_4(f) = \frac{1}{q-1} \left( 1 - \iint f^q(x_1, x_2) dx_1 dx_2 \right), \quad q \geq 0 \text{ et } q \neq 1.$$

En prenant la dérivée par rapport à  $f$  dans l'équation (A.55), les autres fonctions  $\lambda_1(x_1), \lambda_2(x_2), f_1(x_1), f_2(x_2)$  **sont fixées** et on calcule la “dérivée”(en supposant son existence ) suivante

$$\frac{\partial \mathcal{L}_4(f)}{\partial f} = 0. \quad (\text{A.56})$$

$$\frac{d}{df} [J_4(f)] = \frac{1}{q-1} \left( - \iint \frac{d}{df} [f^q(x_1, x_2)] dx_1 dx_2 \right) \quad (\text{A.57})$$

$$= -\frac{1}{q-1} \iint q f^{q-1}(x_1, x_2) dx_1 dx_2 \quad (\text{A.58})$$

$$= \iint \left( -\frac{q}{q-1} \right) f^{q-1}(x_1, x_2) dx_1 dx_2. \quad (\text{A.59})$$

De même pour la dérivée par rapport à  $f$  du terme suivant dans l'équation (A.55), on a :

$$\frac{d}{df} \left[ \lambda_0 \left( 1 - \iint f(x_1, x_2) dx_1 dx_2 \right) \right] = -\lambda_0 \left( \iint \frac{d}{df} [f(x_1, x_2)] dx_1 dx_2 \right) \quad (\text{A.60})$$

$$= -\lambda_0 \left( \iint dx_1 dx_2 \right) \quad (\text{A.61})$$

$$\begin{aligned}
 \frac{d}{df} \left[ \int \lambda_1(x_1) \left( f_1(x_1) - \int f(x_1, x_2) dx_2 \right) dx_1 \right] &= \int \lambda_1(x_1) \left( - \int \frac{d}{df} [f(x_1, x_2)] dx_1 dx_2 \right) \\
 &= \int \lambda_1(x_1) \left( - \int dx_1 dx_2 \right) \\
 &= \int \int (-\lambda_1(x_1)) dx_1 dx_2 \quad (A.62)
 \end{aligned}$$

De même pour le dernier terme dans l'équation (A.55), on a :

$$\begin{aligned}
 \frac{d}{df} \left[ \int \lambda_2(x_2) \left( f_2(x_2) - \int f(x_1, x_2) dx_2 \right) dx_1 \right] &= \int \lambda_2(x_2) \left( - \int \frac{d}{df} [f(x_1, x_2)] dx_1 dx_2 \right) \\
 &= \int \lambda_2(x_2) \left( - \int dx_1 dx_2 \right) \\
 &= \int \int (-\lambda_2(x_2)) dx_1 dx_2. \quad (A.63)
 \end{aligned}$$

En utilisant les équations (A.59), (A.61), (A.62) et (A.63) l'expression dans l'équation (A.56) devient

$$\begin{aligned}
 \frac{\partial \mathcal{L}_4(f)}{\partial f} &= \iint \left( -\frac{q}{q-1} \right) f^{q-1}(x_1, x_2) dx_1 dx_2 - \lambda_0 \left( \iint dx_1 dx_2 \right) + \\
 &\quad \int \int (-\lambda_1(x_1)) dx_1 dx_2 + \int \int (-\lambda_2(x_2)) dx_1 dx_2. \quad (A.64)
 \end{aligned}$$

Pour question de rigueur mathématique (par exemple l'existence de la somme des intégrables des fonctions inconnues  $\lambda_1(x_1), \lambda_2(x_2)$  dans les différentes intégrales dont les bornes sont inconnues), j'ai imposé par exemple que  $0 \leq x_1 \leq 1$  et  $0 \leq x_2 \leq 1$ , ce qui permet d'avoir dans ce cas  $\iint \equiv \int_0^1 \int_0^1$  et par conséquent l'équation (A.64) devient :

$$\begin{aligned}
 \frac{\partial \mathcal{L}_4(f)}{\partial f} &= \int_0^1 \int_0^1 \left( -\frac{q}{q-1} \right) f^{q-1}(x_1, x_2) dx_1 dx_2 - \lambda_0 \left( \int_0^1 \int_0^1 dx_1 dx_2 \right) + \\
 &\quad \int_0^1 \int_0^1 (-\lambda_1(x_1)) dx_1 dx_2 + \int_0^1 \int_0^1 (-\lambda_2(x_2)) dx_1 dx_2. \quad (A.65)
 \end{aligned}$$

On obtient alors :

$$\frac{\partial \mathcal{L}_4(f)}{\partial f} = \int_0^1 \int_0^1 \left[ -\frac{q}{q-1} f^{q-1}(x_1, x_2) - \lambda_0 - \lambda_1(x_1) - \lambda_2(x_2) \right] dx_1 dx_2 = 0$$

En supposant que la fonction  $-\frac{q}{q-1} f^{q-1}(x_1, x_2) - \lambda_0 - \lambda_1(x_1) - \lambda_2(x_2)$  est continue sur  $[0, 1]^2$ , si son intégrale est nulle alors elle est nulle.



## ANNEXE A. SUR LA PREUVE DES DENSITÉS DE PROBABILITÉ MAXIMISANT DES ENTROPIES

---

On a alors

$$-\frac{q}{q-1}f^{q-1}(x_1, x_2) - \lambda_0 - \lambda_1(x_1) - \lambda_2(x_2) = 0 \quad (\text{A.66})$$

$$-\frac{q}{q-1}f^{q-1}(x_1, x_2) = \lambda_0 + \lambda_1(x_1) + \lambda_2(x_2) \quad (\text{A.67})$$

$$f^{q-1}(x_1, x_2) = -\frac{q-1}{q}(\lambda_0 + \lambda_1(x_1) + \lambda_2(x_2))$$

$$f^{q-1}(x_1, x_2) = \frac{1-q}{q}(\lambda_0 + \lambda_1(x_1) + \lambda_2(x_2))$$

$$f(x_1, x_2) = \left[ \frac{1-q}{q}(\lambda_0 + \lambda_1(x_1) + \lambda_2(x_2)) \right]^{\frac{1}{q-1}}$$

Pour le cas particulier où l'indexe de l'entropie de Tsallis  $q = 2$ , on a

$$f(x_1, x_2) = -\frac{1}{2}(\lambda_0 + \lambda_1(x_1) + \lambda_2(x_2)) \quad (\text{A.68})$$

La démonstration précédente est similaire pour l'entropie de Rényi et pour  $q = 2$ , on obtient

$$f(x_1, x_2) = -\frac{1}{2}m(\lambda_0 + \lambda_1(x_1) + \lambda_2(x_2)), \quad (\text{A.69})$$

avec  $m = \int_0^1 \int_0^1 f^2(x_1, x_2) dx_1 dx_2$ .

De plus si on impose  $m = 1$ , les deux équations (A.68) et (A.69) coïncident.

# Diffusion des résultats obtenus



# Communications avec actes dans des conférences internationales

## Papier I

### **Copula and tomography :**

Paper number 107 accepted and presented at VISAPP 2009 -  
International Conference on Computer Vision Theory and  
Applications (5-8 February 2009, Lisboa, Portugal)



## Papier II

**Lien entre la notion de copule et la tomographie :**  
Peer review paper in Proceedings of GRETSI 2009 : GRETSI  
Symposium on Signal and Image Processing , 8-11 September  
2009 Dijon, France.



---

Article dans une revue internationale avec comité de lecture

---

## Papier III

### **Link between copula and tomography :**

Peer review article published in Pattern Recognition Letters-  
Elsevier (An official publication of the International  
Association for Pattern Recognition), Volume 31, Issue 14,  
Pages 2258-2264





## Papier III

### **Maximum Entropies Copulas :**

Peer review Paper in the Conference Proceedings of the  
American Institute of Physics (AIP), Subseries : Mathematical  
and Statistical Physics , Volume 1305, Issue 1, Pages 329-336 .



# Papier IV

**New Copulas obtained by maximizing Tsallis or Rényi  
Entropies :**

Peer review paper in the Conference Proceedings of the  
American Institute of Physics (AIP), Subseries : Mathematical  
and Statistical Physics , Volume 0000, Issue 0, Pages 000-000 .



# Bibliographie

- [1] J. Radon. Über die bestimmung von funktionen durch ihre integralwerte längs gewisser mannigfaltigkeiten. *Ber. Verh. Säch. Akad. Wiss. Leipzig, Math. Nat. Kl.*, 69 :262–277, 1917.
- [2] R.N. Bracewell. Strip integration in radio astronomy. *Aust. J. Phys.*, 9 :198–217, 1956.
- [3] S.R. Deans. The Radon transform and some of its applications. *A Wiley-Interscience Publication, New York*, 1983.
- [4] M. Fréchet. *Les Probabilités associées à un système d'évènements compatibles et dépendants, Partie I*, volume 859. Paris : Hermann, 1940.
- [5] A. Sklar. Fonctions de répartition à  $n$  dimensions et leurs marges. *Publ. Inst. Statist. Univ. Paris 8*, pages 229–231, 1959.
- [6] J. Hadamard. Sur les problèmes aux dérivées partielles et leur signification physique. *Princeton University Bulletin*, 13 :49–52, 1902.
- [7] W. Hoeffding. Scale-Invariant Correlation Theory. *The Collected Works of Wassily Hoeffding. New York : Springer-Verlag*, pages 57–107, 1940.
- [8] W. Hoeffding. Scale-invariant correlation measures for discontinuous distributions. *The Collected Works of Wassily Hoeffding. New York : Springer-Verlag*, pages 109–133, 1941.
- [9] Roger B. Nelsen. *An Introduction to Copulas*. Springer Verlag, 2nd edition, 2006.
- [10] H. Cramer and H. Wold. Some theorems on distribution functions. *Journal of the London Mathematical Society*, 1(4) :290–294, 1936.
- [11] A. Rényi. On projections of probability distributions. *Acta Mathematica Hungarica*, 3(3) :131–142, 1952.
- [12] W.M. Gilbert. Projections of probability distributions. *Acta Mathematica Hungarica*, 6(1) :195–198, 1955.
- [13] G.G. Lorentz. A problem of plane measure. *American Journal of Mathematics*, 71(2) :417–426, 1949.
- [14] H.J. Ryser. Combinatorial properties of matrices of zeros and ones. *Canad. J. Math.*, 9 :371–377, 1957.
- [15] H.J. Ryser. Matrices of zeros and ones. *Bull. Amer. Math. Soc*, 66 :442–464, 1960.
- [16] G.T. Herman and A. Kuba. *Advances in discrete tomography and its applications*. Birkhäuser, 2007.
- [17] W. F. Darsow, B. Nguyen, and E. T. Olsen. Copulas and Markov processes. *Illinois J. Math.*, 36 :600–642, 1992.
- [18] O. Darrigol. The origins of the entropy concept. In *Poincaré Seminar 2003 : Bose-Einstein condensation-entropy*, volume 38, page 101. Birkhauser, 2004.
- [19] C.E. Shannon. A mathematical theory of communication. *Bell System Technical Journal*, 27 :432–379, 1948.
- [20] A. Rényi. On measures of dependence. *Acta Mathematica Hungarica*, 10(3) :441–451, 1959.

- 
- [21] A. Rényi. On measures of entropy and information. In *Proceedings of the 4th Berkeley Symposium on Mathematical Statistics and Probability*, volume 1, pages 547–561, 1961.
- [22] J. Havrda and F. Charvát. Quantification method of classification processes. *Kybernetika*, 3(1) :30–35, 1967.
- [23] J.P. Burg. *Maximum entropy spectral analysis*. PhD thesis, Stanford University, 1975.
- [24] C. Tsallis. Possible generalization of Boltzmann-Gibbs statistics. *Journal of statistical physics*, 52(1) :479–487, 1988.
- [25] E.T. Jaynes. Information Theory and Statistical Mechanics. *Physical Review*, 106(4) :620–630, 1957.
- [26] E.T. Jaynes. Information theory and statistical mechanics. II. *Physical Review*, 108(2) :171–190, 1957.
- [27] A.N. Tikhonov, V.I.A. Arsenin, and F. John. *Solutions of ill posed problems*. Winston Washington, DC :, 1977.
- [28] E. Serge. Nuclei and particles. An introduction to nuclear and subnuclear physics. 1977.
- [29] A. Kak and M. Slaney. *Principles of Computerized Tomographic Imaging*. Society of Industrial and Applied Mathematics, 1988.
- [30] A. Tarantola. *Inverse Problem Theory and Methods for Model Parameter Estimation*. SIAM : Society for Industrial and Applied Mathematics, 1 edition, December 2004.
- [31] J. Idier. *Bayesian Approach to Inverse Problems*. Wiley-ISTE, 1 edition, june 2008.
- [32] K.F. Koral, N.H. Clinthorne, W.L. Rogers, and J.W. Keyes Jr. Feasibility of sharpening limited-angle tomography by including an orthogonal set of projections. *Nuclear Instruments and Methods in Physics Research*, 193(1-2) :223–227, 1982.
- [33] N.V. Denisova. Two-view tomography. *Journal of Physics D : Applied Physics*, 33 :313, 2000.
- [34] J. Radon. On the determination of functions from their integral values along certain manifolds. *IEEE transactions on medical imaging*, 5(4) :170–176, 1986.
- [35] F. Natterer. *The mathematics of computerized tomography*. Society for Industrial Mathematics, 2001.
- [36] A. Markoe. *Analytic Tomography*. Cambridge University Press, 2006.
- [37] S. Helgason. *The radon transform*, volume 5. Birkhauser, 1999.
- [38] S. Helgason. *Geometric analysis on symmetric spaces*. American Mathematical Society, 2008.
- [39] H. Joe. *Multivariate Models and Dependence Concepts*. London : Chapman & Hall, 1997.
- [40] C.E. Priebe. Adaptive mixtures. *Journal of the American Statistical Association*, pages 796–806, 1994.
- [41] J. Havil. *Gamma : Exploring Euler’s Constant*. Princeton University Press, 2003.
- [42] B. Schweizer and Abe Sklar. *Probabilistic Metric Spaces*. North Holland New York, 1983.

- 
- [43] Ali M. M., Mikhail N. N., and Haq M. S. A class of bivariate distributions including the bivariate logistics. *J. Multivariate Anal.*, 8 :405–412, 1978.
- [44] M. Fréchet. Sur les tableaux de corrélation dont les marges sont données. *Ann. Univ. Lyon, III. Sér., Sect. A*, 14 :53–77, 1951.
- [45] A. Sklar. Random Variables, Distribution Functions, and Copulas : A Personal Look Backward and Forward. *Lecture Notes-Monograph Series*, pages 1–14, 1996.
- [46] J. Yan. Enjoy the joy of copulas : With a package copula. *Journal of Statistical Software*, 21(i04), 2007.
- [47] A.K. Nikoloulopoulos and K. Dimitris. Copula model evaluation based on parametric bootstrap. *Computational Statistics & Data Analysis*, 52 :3342–3353, 2008.
- [48] H. Jeffreys. An invariant form for the prior probability in estimation problems. *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences*, pages 453–461, 1946.
- [49] D. Huart, G. Évin, and A.-C. Favre. Bayesian copula selection. *Computational Statistics & Data Analysis*, 51 :809–822, 2006.
- [50] C. Genest and J. MacKay. The joy of copulas : Bivariate distributions with uniform marginals. *The American Statistician*, 40(4) :280–283, 1986.
- [51] V. Durrleman, A. Nikeghbali, and T. Roncalli. Which copula is the right one? *Working document, Groupe de Recherche Opérationnelle, Crédit Lyonnais*, 25 August 2000.
- [52] W.C.M. Kallenberg. Modelling dependence. *Insurance : Mathematics and Economics*, 42 :127–146, 2008.
- [53] T. Mikosch. Copulas : Tales and facts. *Extremes*, 9(1) :3–20, 2006.
- [54] C. Genest and B. Rémillard. Discussion of "copulas : Tales and facts", by Thomas Mikosch. *Extremes*, 9(1) :27–36, 2006.
- [55] P. Embrechts. Discussion of "copulas : Tales and facts", by Thomas Mikosch. *Extremes*, 9(1) :45–47, 2006.
- [56] H. Joe. Discussion of "copulas : Tales and facts", by Thomas Mikosch. *Extremes*, 9(1) :37–41, 2006.
- [57] C.G. de Vries and C. Zhou. Discussion of "copulas : Tales and facts", by Thomas Mikosch. *Extremes*, 9(1) :23–25, 2006.
- [58] P. Embrechts. Copulas : A personal view. *Journal of Risk and Insurance*, 76(3) :639–650, 2009.
- [59] C. Genest and A.-C. Favre. Everything you always wanted to know about copula modeling but were afraid to ask. *Journal of Hydrologic Engineering*, 12 :347–368, 2007.
- [60] L. Rüschendorf. Construction of multivariate distributions with given marginals. *Annals of the Institute of Statistical Mathematics*, 37(1) :225–233, 1985.
- [61] C.D. Lai and M. Xie. A new family of positive quadrant dependent bivariate distributions. *Statistics & probability letters*, 46(4) :359–364, 2000.
- [62] J.N. Kapur and H.K. Kesavan. *Entropy optimization principles with applications*. Academic Press, 1992.
- [63] E.T. Jaynes. Gibbs vs Boltzmann entropies. *American Journal of Physics*, 33 :391, 1965.
-



- 
- [64] C. Tsallis. *Introduction to Nonextensive Statistical Mechanics : Approaching a Complex World*. Springer Verlag, 2009.
- [65] V. Strassen. The existence of probability measures with given marginals. *The Annals of Mathematical Statistics*, 36(2) :423–439, 1965.
- [66] S. Kullback. Probability densities with given marginals. *The Annals of Mathematical Statistics*, pages 1236–1243, 1968.
- [67] I. Csiszar. I-divergence geometry of probability distributions and minimization problems. *Ann. Probab*, 3(1) :146–158, 1975.
- [68] J.M. Borwein, A.S. Lewis, and R.D. Nussbaum. Entropy minimization, DAD problems, and doubly stochastic kernels. *Journal of Functional Analysis*, 123(2) :264–307, 1994.
- [69] A.M.H. Meeuwissen and T. Bedford. Minimally informative distributions with given rank correlation for use in uncertainty analysis. *Journal of Statistical Computation and Simulation*, 57(1) :143–174, 1997.
- [70] E. Pasha and S. Mansoury. Determination of Maximum Entropy Multivariate Probability Distribution under some Constraints. *Applied Mathematical Sciences*, 2(57) :2843–2849, 2008.
- [71] A.W. Marshall and I. Olkin. *Life distributions : Structure of nonparametric, semi-parametric, and parametric families*. Springer Verlag, 2007.
- [72] Y. Censor, A.R. De Pierro, and A.N. Iusem. Optimization of Burg’s entropy over linear constraints. *Applied Numerical Mathematics*, 7(2) :151–165, 1991.
- [73] D. Ellerman. Counting distinctions : on the conceptual foundations of Shannon’s information theory. *Synthese*, 168(1) :119–149, 2009.
- [74] L. Nivanen, A. Le Mehaute, and QA Wang. Generalized algebra within a nonextensive statistics. *Reports on Mathematical Physics*, 52(3) :437–444, 2003.
- [75] P. Kumaraswamy. A generalized probability density function for double-bounded random processes. *Journal of Hydrology*, 46(1-2) :79–88, 1980.
- [76] M.C. Jones. Kumaraswamy’s distribution : A Beta-type distribution with some tractability advantages. *Statistical Methodology*, 6(1) :70–81, 2009.
- [77] J.R. Van Dorp and S. Kotz. The Standard Two-Sided Power Distribution and its Properties : With application in financial engineering. *The American Statistician*, 56(2) :90–99, 2002.
- [78] J.A. Rodríguez-Lallena and M. Úbeda-Flores. A new class of bivariate copulas. *Statistics & probability letters*, 66(3) :315–325, 2004.
- [79] W. F. Darsow, B. Nguyen, and E. T. Olsen. Copulas and markov processes. *Illinois J. Math.*, 36 :600–642, 1992.
- [80] G. Mayor, J. Suner, and J. Torrens. Copula-like operations on finite settings. *IEEE Transactions on fuzzy systems*, 13(4) :468–477, 2005.
- [81] A. Kolesárová, R. Mesiar, J. Mordelová, and C. Sempi. Discrete copulas. *IEEE Transaction on fuzzy systems*, 14(5), 2006.
- [82] G.T. Herman and A. Kuba. *Discrete Tomography : Foundations, Algorithms, and Applications*. Birkhäuser, 1999.
- [83] G. Birkhoff. Tres observaciones sobre el algebra lineal. Univ. Nac. *Tucuman Revista A*, 5 :147–151, 1946.
-

- 
- [84] W. Glunt, T.L. Hayden, and R. Reams. The nearest generalized doubly stochastic matrix to a real matrix with the same first and second moments. *Comp. Appl. Math*, 27(2), 2008.
- [85] H.J. Ryser and J. Cofman. *Combinatorial Mathematics*. Mathematical Association of America, 1963.
- [86] R.L. Plackett. A class of bivariate distributions. *Journal of the American Statistical Association*, 60(310) :516–522, 1965.
- [87] Geert Molenberghs and Emmanuel Lesaffre. Marginal modeling of correlated ordinal data using a multivariate plackett distribution. *Journal of the American Statistical Association*, 89(426) :633–644, 1994.
- [88] RB Nelsen, JJ Quesada-Molina, and JA Rodríguez-Lallena. Bivariate copulas with cubic sections. *Journal of Nonparametric Statistics*, 7(3) :205–220, 1997.
- [89] Dolores A. Conway. Multivariate distributions with specified marginals. Technical Report 145, Stanford University, Dept. of Statistics, August 1979.
- [90] P.J. Szablowski. q-Gaussian distributions : Simplifications and simulations. *Journal of Probability and Statistics*, 2009.
- [91] S. Nadarajah and S. Kotz. On the q-type distributions. *Physica A : Statistical Mechanics and its Applications*, 377(2) :465–468, 2007.
- [92] J.F. Bercher. Escort entropies and divergences and related canonical distribution. *Physics Letters A*, 2011.
- [93] A. Cichocki, S. Cruces, and S. Amari. Generalized alpha-beta divergences and their application to robust nonnegative matrix factorization. *Entropy*, 13(1) :134–170, 2011.
- [94] C. Beck and F. Schlögl. *Thermodynamics of chaotic systems : an introduction*, volume 4. Cambridge University Press, 1995.
- [95] S. Umarov, C. Tsallis, and S. Steinberg. On a q-central limit theorem consistent with nonextensive statistical mechanics. *Milan journal of mathematics*, 76(1) :307–328, 2008.
- [96] HJ Hilhorst. Central limit theorems for correlated variables : some critical remarks. *Brazilian Journal of Physics*, 39(2A) :371–379, 2009.