Supermodular comparison inequalities in option pricing and information inequalities

Kizildemir, Bunyamin

2017

Kizildemir, B. (2017). Supermodular comparison inequalities in option pricing and information inequalities. Doctoral thesis, Nanyang Technological University, Singapore.

<http://hdl.handle.net/10356/69531>

<https://doi.org/10.32657/10356/69531>

NANYANG
TECHNOLOGICAL
UNIVERSITY

**Supermodular Comparison Inequalities In
Option Pricing and Information
Inequalities**

BÜNYAMİN KIZILDEMİR

**SCHOOL OF PHYSICAL AND MATHEMATICAL
SCIENCES**

2017

Supermodular Comparison Inequalities In Option Pricing and Information Inequalities

BÜNYAMİN KIZILDEMİR

School of Physical and Mathematical Sciences

A thesis submitted to the Nanyang Technological University
in partial fulfilment of the requirement for the degree of
Doctor of Philosophy

2017

Abstract

This thesis is devoted to find necessary and sufficient conditions for supermodular ordering of Poisson random vectors, to bound option prices for multidimensional Ornstein-Uhlenbeck process and multidimensional jump-diffusion model, and to derive inequalities for information theory measures. First, we find necessary and sufficient conditions for supermodular ordering of Poisson random vectors of a certain structure and then we derive the same necessary and sufficient conditions for supermodular ordering of Binomial random vectors, and in the limit case we recover supermodular ordering of Poisson random vectors and existing results in the literature for Gaussian random vectors. Second, by using stochastic and supermodular ordering we derive sufficient conditions for bounding option prices for multidimensional Ornstein-Uhlenbeck process and multidimensional jump-diffusion model. Lastly, we calculate the entropy and mutual information of multidimensional Brownian motion and Poisson process by using Itô calculus and then we derive some comparison results.

Acknowledgement

PhD life is a long journey with full of excitements, stress, ups and downs. In this long journey having a wise and helpful supervisor is like having a torch in the dark. Many thanks to my supervisor, Prof Nicolas Privault, without his suggestions and help I couldn't come to this point. Maybe the most important thing that I learned from him is how to do research. For me, this is the fruit of PhD, the essence of PhD. Again I want to thank to my golden hearted supervisor Prof Nicolas Privault.

I want to thank to Prof Chee Yeow Meng and Prof Wu Guohua for their help and guide during my PhD years. In my language, Turkish, we have a saying "if someone teach me a word, I will be his/her slave for 40 years", for me this sentence best describe how value teachers are. Therefore, I want to thank Prof Xiang Liming who taught me 'statistics' and helped me to conduct tutorial sessions for her undergraduate classes, Prof Wang Li-Lian who taught me 'partial differential equations' and accepted to become a member of my thesis advisory committee, Prof Wu Guohua who taught me 'continuous methods' and guide me whenever I need, Prof Pan Guangming from whom I took graduate seminar course and who has always supported me with his friendly suggestions, and of course, special thanks to my teacher and supervisor Prof Nicolas Privault. Also I want to point out my thanks to Prof Adrian Roellin, from Department of Statistics and Applied Probability of the National University of Singapore, who accepted to be a member of my thesis advisory committee and helped me whenever I need. Special thanks to Prof Lam Khin Yong who helped me and to my Turkish friends whenever we need.

Without encourage and support of my family I think I couldn't achieve success in my life. Therefore my special thanks to my father, mother and to my lovely siblings. My classmates, research lab friends deserve many thanks, especially I want to thank to Dr Liu Yue and Dr Ian Flint who were always with me in my ups and downs.

Contents

Abstract	i
Acknowledgement	iii
Acknowledgement	v
Introduction	1
1 Background	7
1.1 Stochastic ordering of random vectors	7
1.2 Binomial and Poisson random vectors	13
1.3 Lévy Measures	14
1.4 Möbius Inversion formula	16
1.5 Information theory measures	18
1.6 Log-Sobolev inequality on the Poisson space	22
I Stochastic and supermodular ordering	27
2 Supermodular ordering of Poisson arrays	29
2.1 Introduction	30
2.2 Poisson random vectors	32

2.3	Triangular Poisson arrays	34
2.4	Supermodular Poisson arrays	38
3	Supermodular ordering by tree-based correlations	49
3.1	Introduction	50
3.2	Tree-based correlation structures	52
3.3	Binomial random vectors	58
3.4	Poisson random vectors	67
3.5	Some counter examples	75
4	Bounds for option prices by supermodular ordering	85
4.1	Introduction	85
4.2	Preservation of supermodularity for multidimensional Ornstein-Uhlenbeck process	88
4.3	Bounds of option prices for multidimensional Ornstein-Uhlenbeck process	90
4.4	Preservation of supermodularity for multidimensional jump-diffusion model	96
4.5	Locally supermodularity preserving operators	98
II	Entropy and mutual information	105
5	Information theory on the Wiener space	107
5.1	Introduction	107
5.2	Comparison of mutual information for Gaussian random vectors	108

5.3	Differential entropy and mutual information of Brownian motion	111
5.4	Differential entropies for multidimensional bifractional Brownian motion	116
5.5	Log-Sobolev type inequalities for Gaussian densities	120
6	Information theory on the Poisson space	125
6.1	Introduction	125
6.2	Entropy of Poisson random variables	126
6.3	Entropy of Poisson process	128
6.4	Entropy of the compound Poisson and the Gamma processes	130
6.5	Log-Sobolev type inequalities for the relative entropy . . .	134
6.6	Calculations for predictable decompositions	139
	Future Research	145
	Conclusion	149
	Bibliography	156

Introduction

Mathematics is the most
beautiful and most powerful
creation of the human spirit.

Stefan Banach

Stochastic ordering is one of the developing research area of mathematics with potential questions and applications. Alfred Müller, Marco Scarsini, Dietrich Stoyan, Moshe Shaked, J. George Shanthikumar, and Ludger Rüschendorf can be seen as pioneers of this research area of mathematics. They give lots of applications of stochastic ordering in their papers and books these include insurance mathematics, risk management, stochastic networks, financial mathematics, statistical physics and many others. The main idea behind stochastic ordering and comparison is to derive approximation results and bounds for complex stochastic models whose analytic solution is not known. Information entropy was introduced by Claude Shannon [36] and widely used in engineering and in science as a measure of uncertainty.

Supermodular ordering of Gaussian random vectors has been completely characterized in Müller and Scarsini [25] and in Müller and Stoyan [26]. They show that for supermodular ordering of random vectors X and Y which are Gaussian necessary and sufficient conditions are

- (i) X_i and Y_i have same distribution for all $i = 1, \dots, d$, and

(ii) for all $1 \leq i < j \leq d$ we have

$$\text{Cov}(X_i, X_j) \leq \text{Cov}(Y_i, Y_j).$$

Kızıldemir and Privault [18] showed that (i) and (ii) above are also necessary and sufficient conditions for supermodular ordering of Poisson random vectors for a certain structure. Then Kızıldemir and Privault [17] extended previous results for supermodular ordering of Poisson random vectors to a much larger family; their approach cover binomial random vectors and recovers the Gaussian setting of [25] as a consequence.

In Chapter 2, we show that the same necessary and sufficient conditions hold while X and Y are Poisson random vectors of a certain structure. Since any d -dimensional Poisson random vector is based on $2^d - 1$ parameters, therefore (i) and (ii), which are based on $d(d+1)/2$ conditions, cannot characterize their distribution ordering except if $d = 2$. Therefore we consider a particular dependence structure of Poisson arrays depending on $d(d+1)/2$ parameters based on a natural decomposition of their Lévy measures on the vertices of the d -dimensional unit hypercube. We also propose some financial applications of stochastic ordering. We derive sufficient conditions for the supermodular ordering of sums of Gaussian and Poisson random vectors, with application to a jump-diffusion asset price model.

In Chapter 3, we construct a general dependence structure for binomial and Poisson random vectors, under which supermodular ordering can be characterized by the ordering of covariance matrices. This approach extends our previous results to a much larger family of dependence struc-

tures, it covers the setting of binomial random vectors and recovers the Gaussian setting of [25] as a consequence. The construction is based on the arrangement of random variables according to a binary tree, and uses the Möbius inversion.

Stochastic ordering is not only limited to the supermodular functions; there are lots of research about stochastic ordering of convex functions and other functions as well; the book "Comparison Methods for Stochastic Models and Risks" [26] is an essential book for stochastic ordering of random vectors with financial applications. Similarly, applications of stochastic ordering is not only limited with insurance mathematics, risk management, stochastic networks, financial mathematics, and statistical physics: also there are lots of applications of stochastic ordering of random vectors, including in economics and in game theory, the paper of Meyer and Strulovici [21] include many such applications.

Bergenthum and Rüschendorf [5], [4] give sufficient conditions for comparison of European options with respect to martingale pricing measures. They use convex and directionally convex payoff functions for their comparison. Technically they need the "propagation of convexity property" and they establish this property for several further examples of univariate and multivariate processes. In application part, they give comparison results for multivariate diffusion processes, stochastic volatility models, Lévy processes, and diffusion with jumps.

In Chapter 4, we find bounds for multidimensional stochastic processes by using multidimensional Ornstein-Uhlenbeck process and multidimensional jump-diffusion model. Our comparison is with respect to supermodular stochastic ordering and, since the analytic solution of mul-

tidimensional Ornstein-Uhlenbeck process and multidimensional jump-diffusion model are known, technically it is possible to prove "the propagation of supermodularity property". For general multidimensional processes whose analytic solution is not known it is hard to prove this property because of the diffusion part in the process. Ornstein-Uhlenbeck process is one of the famous process that is used to model interest rates, currency exchange rates, and commodity prices stochastically. Lastly, we state a definition for locally supermodularity preserving (LSP) operators and we show that if a model is supermodularity preserving then it is LSP.

Entropy is not only a concept of physics, it is used a lot in almost every branch of science and engineering. It is also used widely in finance and economics: for example, Ormos and Zibriczky [30] use entropy as a risk measure; Stutzer [38] derives Black-Scholes Option Pricing model by using relative entropy; Zhou et al. [44] give applications of the entropy for portfolio selection and asset pricing. The book "elements of information theory" by Cover and Thomas [9] is one of the key book for concepts related to entropy, relative entropy, and mutual information. It is shown in Narasimhan and Bilmes [27] that entropy and mutual information are supermodular functions. Also Cicalese and Vaccaro [8] shows supermodularity and subadditivity Properties of the Entropy on the Majorization Lattice.

In Chapter 5, we state necessary and sufficient conditions for comparison of mutual information for Gaussian random vectors. We calculate the differential entropy and mutual information of Brownian motion, and then by using Itô formula we calculate difference of differential entropies for multidimensional Brownian motion. Then we obtain a more general

formula for the difference of differential entropies by using the Itô-Tanaka formula for d -dimensional bifractional Brownian motion. Nourdin et al. [29] and Besalu et al. [6] provide bounds for Gaussian type densities and we state log-Sobolev type inequalities by using their density function.

In Chapter 6, we derive a connection formula between information entropy and probability entropy for Poisson random variables. Later we derive the same connection formula for Poisson process by using Itô calculus. Moreover we calculate the entropy of the compound Poisson process and the Gamma processes. Clark-Ocone theorem is one of the important theorem of stochastic analysis which is used to hedge financial derivatives. The last two sections of this chapter is motivated by the Clark-Ocone theorem. Wu [41] established a modified logarithmic Sobolev inequality by martingale representation approach and we extend the logarithmic Sobolev inequality of [41] to the relative entropy of two probability measures on the Poisson space.

List of Publications

Kızıldemir, B., and Privault, N. Supermodular ordering of binomial, Poisson and Gaussian random vectors by tree-based correlations.
<http://arxiv.org/abs/1512.08883>. Submitted.

Kızıldemir, B., and Privault, N. Supermodular ordering of Poisson arrays . *Statistics and Probability Letters* 98, (2015), 136 - 143.

Chapter 1

Background

1.1 Stochastic ordering of random vectors

In this background section on stochastic ordering, first, we provide the definitions of stochastic, convex, increasing convex, and supermodular ordering of random vectors and we also state definitions for convex and supermodular functions, and then we list some properties of stochastic orders. Moreover, we state necessary and sufficient conditions for supermodular ordering of Gaussian random vectors. Second, we mention about Binomial and Poisson random vectors. Third, we mention about Lévy processes and then we state a remarkable result: supermodular ordering of Lévy processes is equivalent to supermodular ordering of their Lévy measures. Lastly, we state a formula from number theory, Möbius inversion formula, which makes our proofs easier in Chapter 3.

A d -dimensional random vector $X = (X_1, \dots, X_d)$ is said to be dom-

inated by another random vector $Y = (Y_1, \dots, Y_d)$ if

$$E[\Phi(X)] \leq E[\Phi(Y)], \quad (1.1.1)$$

for all sufficiently integrable $\Phi : \mathbb{R}^d \rightarrow \mathbb{R}$ within a certain class of functions that determines the ordering.

Definition 1.1.1. *i) One writes $X \leq_{\text{st}} Y$, the stochastic ordering, if (1.1.1) holds for all increasing functions.*

ii) One writes $X \leq_{\text{cx}} Y$, the convex ordering, if (1.1.1) holds for all convex functions, i.e. for every function $\Phi : \mathbb{R}^d \rightarrow \mathbb{R}$ such that

$$\Phi(\alpha x + (1 - \alpha)y) \leq \alpha\Phi(x) + (1 - \alpha)\Phi(y), \quad \alpha \in [0, 1], \quad x, y \in \mathbb{R}^d.$$

Equivalently, a twice differentiable function $f : \mathbb{R}^d \rightarrow \mathbb{R}$ is convex if and only if its Hessian matrix

$$H(x) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_d} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_d} \\ \vdots & \vdots & \cdots & \vdots \\ \frac{\partial^2 f}{\partial x_d \partial x_1} & \frac{\partial^2 f}{\partial x_d \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_d^2} \end{pmatrix}$$

is positive semidefinite.

iii) One writes $X \leq_{\text{sl}} Y$, the stop-loss ordering, when (1.1.1) holds for all non-decreasing convex functions.

Definition 1.1.2. *One writes $X \leq_{\text{sm}} Y$, the supermodular ordering, if (1.1.1) holds for all supermodular functions, i.e. for every function $\Phi :$*

$\mathbb{R}^d \longrightarrow \mathbb{R}$ such that

$$\Phi(x) + \Phi(y) \leq \Phi(x \wedge y) + \Phi(x \vee y), \quad x, y \in \mathbb{R}^d,$$

where the maximum $\vee$ and the minimum $\wedge$ are defined with respect to the componentwise order.

Equivalently, a twice differentiable function $f : \mathbb{R}^d \rightarrow \mathbb{R}$ is called supermodular if and only if

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{x}) \geq 0 \quad \text{for all } \mathbf{x} \in \mathbb{R}^d \text{ and all } 1 \leq i < j \leq d.$$

Moreover, a twice differentiable function $f : \mathbb{R}^d \rightarrow \mathbb{R}$ is directionally convex if and only if

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{x}) \geq 0 \quad \text{for all } \mathbf{x} \in \mathbb{R}^d \text{ and all } 1 \leq i \leq j \leq d.$$

Property 1.1.3 (cf. Theorem 3.9.3 [26]). *If $g_1, \dots, g_d : \mathbb{R} \rightarrow \mathbb{R}$ are increasing and f is supermodular then*

$$f(g_1(\cdot), \dots, g_d(\cdot))$$

is supermodular.

Before we state the supermodular ordering is closed under convolution; we want to give the definition of weighted probability density function from Chapter 1 of Shaked and Shanthikumar [35].

Let X be a non-negative random variable with probability density function $\mathbb{P}$ and let $b : S \rightarrow [1, \infty)$ be any weight function such that

$\mathbb{E}[b(X)]$ exists, then we define X^b as the random variable with the so-called weighted probability density function $\mathbb{P}_b$ which is given by

$$\mathbb{P}_b(x) = \frac{b(x)\mathbb{P}(x)}{\mathbb{E}[b(X)]}, \quad x \geq 0.$$

The following theorem states that supermodular ordering is closed under convolution.

Theorem 1.1.4 (cf. Theorem 3.9.14-(C) [26]). *Let S be a Polish vector space (i.e. a topological space which admits a separable and complete metric) with a closed order (i.e. if $x_j \preceq y_j$ for all j and $x_j \rightarrow x$, $y_j \rightarrow y$ then $x \preceq y$) and let $b : S \rightarrow [1, \infty)$ be any weight function. Let $\leq_{\text{sm}}$ be supermodular order on $\mathbb{P}_b$. Then if $P_1 \leq_{\text{sm}} P_2$ implies*

$$P_1 \star Q \leq_{\text{sm}} P_2 \star Q$$

for all probability measures Q .

The following theorem states that supermodular ordering is invariant with respect to increasing transforms.

Theorem 1.1.5 (cf. Theorem 3.9.8-(P8) [26]). *The supermodular order is invariant with respect to increasing transforms. That is $(X_1, \dots, X_d) \leq_{\text{sm}} (X'_1, \dots, X'_d)$ implies*

$$(g_1(X_1), \dots, g_d(X_d)) \leq_{\text{sm}} (g_1(X'_1), \dots, g_d(X'_d))$$

for all increasing functions $g_1, \dots, g_d$.

In the following theorem, necessary conditions for supermodular or-

dering of random vectors are given.

Theorem 1.1.6 (cf. Theorem 3.9.5 [26]). *The supermodular order has the following properties:*

- a) If $\mathbf{X} \leq_{\text{sm}} \mathbf{Y}$ then $\mathbf{X}$ and $\mathbf{Y}$ have the same marginals.
- b) If $\mathbf{X} \leq_{\text{sm}} \mathbf{Y}$ then $\text{Cov}(X_i, X_j) \leq \text{Cov}(Y_i, Y_j)$ for all $1 \leq i < j \leq d$.

A random vector $\mathbf{X}$ is said to have a normal distribution $N(\mu, \Sigma)$ with expectation $\mu = \mathbb{E}[\mathbf{X}]$ and covariance matrix $\Sigma = \mathbb{E}[(\mathbf{X} - \mu)(\mathbf{X} - \mu)^T]$ if it has the characteristic function

$$\Psi_{\mathbf{X}}(\mathbf{t}) = \mathbb{E}[e^{i\mathbf{t}^T \mathbf{X}}] = \exp \left(i\mathbf{t}^T \mu - \frac{1}{2} \mathbf{t}^T \Sigma \mathbf{t} \right).$$

If Σ is positive definite and hence invertible then $\mathbf{X}$ has the density

$$\phi_{\mathbf{X}}(\mathbf{x}) = \frac{1}{\sqrt{(2\pi)^d \det(\Sigma)}} \exp \left(-\frac{1}{2} (\mathbf{x} - \mu)^T \Sigma^{-1} (\mathbf{x} - \mu) \right).$$

If $f : \mathbb{R}^d \rightarrow \mathbb{R}$ is twice continuously differentiable, then as usual

$$\nabla f(\mathbf{x}) = \left(\frac{\partial f}{\partial x_i}(\mathbf{x}) \right)_{i=1}^d \quad \text{and} \quad \mathbf{H}_f(\mathbf{x}) = \left(\frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{x}) \right)_{i,j=1}^d$$

are the gradient and the Hessian matrix of f .

In the following theorem, necessary and sufficient conditions for supermodular ordering of Gaussian random vectors are given.

Theorem 1.1.7 (cf. Theorem 4.2 [25] and Theorem 3.13.5 [26]). *Let $\mathbf{X} \sim N(\mu, \Sigma)$ and $\mathbf{Y} \sim N(\mu', \Sigma')$ with $\Sigma = (\sigma_{ij})$ and $\Sigma' = (\sigma'_{ij})$. Then the following conditions are equivalent.*

- (i) $\mathbf{X} \leq_{\text{sm}} \mathbf{Y}$;

(ii) $\mathbf{X}$ and $\mathbf{Y}$ have the same marginals and $\sigma_{ij} \leq \sigma'_{ij}$.

In the following theorem it is shown that supermodularity of random vectors implies stop-loss ordering of marginal sums.

Theorem 1.1.8 (cf. Theorem 3.1 Müller [23]). *Let $\mathbf{X} = (X_1, \dots, X_d)$ and $\mathbf{X}' = (X'_1, \dots, X'_d)$ be random vectors with $\mathbf{X} \leq_{\text{sm}} \mathbf{X}'$ and let*

$$S = \sum_{i=1}^d X_i \quad \text{and} \quad S' = \sum_{i=1}^d X'_i.$$

Then $S \leq_{\text{sl}} S'$.

Consider multivariate distributions with n variables and identical, finite support. The i^{th} variable takes values in $\mathcal{L}_i$ which is a totally ordered set with m_i elements. The Cartesian product $\times_i \mathcal{L}_i$ is denoted $\mathcal{L}$ and is endowed with the usual partial order: $x \leq y$ if and only if $x_i \leq y_i$ for all $i \in \mathcal{N} \equiv \{1, \dots, n\}$.

For any $x, y \in \mathcal{L}$, $x \vee y$ denotes the component-wise maximum of x and y , i.e., the element of $\mathcal{L}$ such that $(x \vee y)_i = \max\{x_i, y_i\} \in \mathcal{L}_i$ for all $i \in \mathcal{N}$. Let $x \wedge y$ similarly denote the component-wise minimum of x, y . A function w is supermodular (on $\mathcal{L}$) if $w(x \vee y) + w(x \wedge y) \geq w(x) + w(y)$ for all $x, y \in \mathcal{L}$.

For any $x \in \mathcal{L}$ such that $x + e_i + e_j \in \mathcal{L}$, let $t_{i,j}^x$ denote the function defined on $\mathcal{L}$ by

$$t_{i,j}^x(x) = t_{i,j}^x(x + e_i + e_j) = 1, \quad t_{i,j}^x(x + e_i) = t_{i,j}^x(x + e_j) = -1,$$

and $t_{i,j}^x(y) = 0$ for all other $y \in \mathcal{L}$. We call $t_{i,j}^x$ an elementary transformation on $\mathcal{L}$, and let $\mathcal{T}$ be the countable set of all elementary transformations.

tions.

Then we have the following theorem for characterization of supermodularity:

Theorem 1.1.9 (cf. Theorem 1 [21]). *$f \leq_{\text{sm}} g$ if and only if there exist nonnegative coefficients $\{\alpha_t\}_{t \in \mathcal{T}}$ such that, with f, g , and t seen as vectors of $\mathbb{R}^d$,*

$$g = f + \sum_{t \in \mathcal{T}} \alpha_t t. \quad (1.1.2)$$

1.2 Binomial and Poisson random vectors

Let $\mathbf{X} = (X_1, X_2, \dots, X_n)$ be n -dimensional random vector, $\mathbf{i} = (i_1, i_2, \dots, i_n)$ be n -dimensional vector with 0, 1 components, and $\mathbf{k} = (k_1, k_2, \dots, k_n)$ be n -dimensional vector with nonnegative integer components. Let $\{\alpha_{\mathbf{i}}, \mathbf{i} = (i_1, i_2, \dots, i_n)\}$ be a sequence of nonnegative integers with $\alpha_{i_1} + \dots + \alpha_{i_n} = N$ and let $\mathbb{P}(\mathbf{X} = \mathbf{i}) = p_{\mathbf{i}}$ be the probability mass function of multivariate Bernoulli distribution with $p_{\mathbf{i}} \geq 0$ and $\sum_{\mathbf{i}} p_{\mathbf{i}} = 1$.

Multivariate binomial distribution, $B(N, p_{\mathbf{i}})$, is defined by

$$P(\mathbf{X} = \mathbf{k}) = \sum_{\sum \alpha_{\mathbf{i}} i_j = k_j} \frac{N!}{\prod_{\mathbf{i}} \alpha_{\mathbf{i}}!} \prod_{\mathbf{i}} p_{\mathbf{i}}^{\alpha_{\mathbf{i}}}.$$

Covariance matrix of $B(N, p_{\mathbf{i}})$ is given by

$$\text{Cov}(X_j, X_k) = N \left(\sum_{i_j=i_k=0} p_{\mathbf{i}} \sum_{i_j=i_k=1} p_{\mathbf{i}} - \sum_{i_j=1, i_k=0} p_{\mathbf{i}} \sum_{i_j=0, i_k=1} p_{\mathbf{i}} \right)$$

and

$$\text{Var}(X_j) = N \left(\sum_{i_j=0} p_{\mathbf{i}} \sum_{i_j=1} p_{\mathbf{i}} \right).$$

Multivariate Poisson distribution is a limiting distribution of $B(N, p_{\mathbf{i}})$ as $N \rightarrow \infty$ under the condition of $Np_{\mathbf{i}} = \lambda_{\mathbf{i}}$ where $\lambda_{\mathbf{i}}$ is a non-negative fixed parameter. If a random vector $\mathbf{X}$ has a binomial distribution $B(N, p_{\mathbf{i}})$ and if we assume $Np_{\mathbf{i}} = \lambda_{\mathbf{i}}$ then we have

$$\lim_{N \rightarrow \infty} P(\mathbf{X} = \mathbf{k}) = \sum_{\sum \alpha_{\mathbf{i}} i_j = k_j} \prod_{\mathbf{i} \neq \mathbf{0}} p(\alpha_{\mathbf{i}}, \lambda_{\mathbf{i}})$$

where $p(\alpha_{\mathbf{i}}, \lambda_{\mathbf{i}})$ is an univariate Poisson probability function.

In the following theorems of Kawamura [16] a decomposition of a Poisson random vector and its covariance matrix is given.

Theorem 1.2.1 (cf. Theorem 3 [16]). *If a random vector $\mathbf{X}$ has the Poisson law, then we have an unique decomposition of the random vector by $X_j = \sum_{i_j=1} X_i$ where X_i ($i \neq 0$) are mutually independent, univariate Poisson variables with parameter λ_i .*

Theorem 1.2.2 (cf. Theorem 4 [16]). *The covariance matrix of the multivariate Poisson distribution is given by*

$$\text{Var}(X_j) = \sum_{i_j=1} \lambda_i, \quad \text{Cov}(X_j, X_k) = \sum_{i_j=i_k=1} \lambda_i.$$

1.3 Lévy Measures

Let $X = (X(t))_{t \geq 0}$ be a d -dimensional Lévy process, i.e. stochastically continuous process with independent and stationary increments.

From the Lévy-Itô-decomposition we know that the distribution of a Lévy process is uniquely determined by a characteristic triplet (γ, A, ν) , where γ is a drift parameter, A is a covariance matrix of a Brownian motion, and ν is the Lévy measure determining the frequency and size of jumps. Then the characteristic function is given as follows:

$$\mathbb{E}[\exp(i\langle u, X_t \rangle)] = \exp \left(t \left(i\langle \gamma, u \rangle - \frac{1}{2} \langle Au, u \rangle + \int_{\mathbb{R}^d} (e^{i\langle u, x \rangle} - 1 - i\langle u, h(x) \rangle) \nu(dx) \right) \right),$$

where the truncation function $h : \mathbb{R}^d \rightarrow \mathbb{R}^d$ is given by $h(u) = (h_1(u), \dots, h_d(u))$ with

$$h_i(u) = u_i \mathbf{1}_{[|u_i| \leq 1]}.$$

As we have to take care of the possible singularity of the measures at zero, we need some technical conditions on the functions to ensure that the occurring integrals are finite. We will denote

$$\mathfrak{B}_0 := \left\{ f : \mathbb{R}^d \rightarrow \mathbb{R} \mid f \text{ is measurable and bounded and } \limsup_{x \rightarrow 0} \frac{|f(x)|}{\|x\|^2} < \infty \right\}.$$

As any Lévy measure ν fulfills $\int (\|x\|^2 \wedge 1) \nu(dx) < \infty$, this implies that for any Lévy measure ν and any $f \in \mathfrak{B}_0$ the integral $\int f d\nu$ is finite. Thus the integrals in the following definition are well-defined.

Definition 1.3.1. *Lévy measures $\nu, \tilde{\nu}$ are said to be comparable with respect to the supermodular order, $\nu \leq_{\text{sm}} \tilde{\nu}$, if for all supermodular $f \in \mathfrak{B}_0$ we have*

$$\int f d\nu \leq \int f d\tilde{\nu}.$$

In the following lemma a characterization of supermodular ordering

of Lévy measures is given.

Lemma 1.3.2 (cf. Lemma 4.2 of Bäuerle et al. [3]). *For Lévy measures $\nu, \tilde{\nu}$ the following conditions are equivalent:*

- (a) $\nu \leq_{\text{sm}} \tilde{\nu}$;
- (b) ν and $\tilde{\nu}$ have the same marginal tail integrals and $\int f d\nu \leq \int f d\tilde{\nu}$ for all increasing supermodular $f \in \mathfrak{B}_0$.

In the following theorem supermodular ordering of the Lévy measures implies supermodular ordering of the corresponding Lévy processes.

Theorem 1.3.3 (cf. Theorem 4.5 [3]). *For two Lévy processes X and $\tilde{X}$ with Lévy measures ν and $\tilde{\nu}$, the following conditions are equivalent:*

- (a) $\nu \leq_{\text{sm}} \tilde{\nu}$;
- (b) $X \leq_{\text{sm}} \tilde{X}$.

1.4 Möbius Inversion formula

Let P be a locally finite partially ordered set. The *incidence algebra* of P is defined as follows. Consider the set of all real-valued functions of two variables $f(x, y)$, defined for x and y ranging over P , and with the property that $f(x, y) = 0$ if $x \not\leq y$. The sum of two such functions f and g , as well as multiplication by scalars, are defined as usual. The product $h = fg$ is defined as follows:

$$h(x, y) = \sum_{x \leq z \leq y} f(x, z)g(z, y).$$

In view of the assumption that P is locally finite, the sum on the right is well-defined. The incidence algebra has an identity element which

we write $\delta(x, y)$, the Kronecker delta. The *zeta function* $\zeta(x, y)$ of the partially ordered set P is the element of the incidence algebra of P such that $\zeta(x, y) = 1$ if $x \leq y$ and $\zeta(x, y) = 0$ otherwise. The function $n(x, y) = \zeta(x, y) - \delta(x, y)$ is called the *incidence function*.

Proposition 1.4.1 (cf. Proposition 1 of Rota [33]). *The zeta function of a locally finite partially ordered set is invertible in the incidence algebra.*

The function μ , inverse to ζ , is called the *Möbius function* of the partially ordered set P .

Proposition 1.4.2 (Möbius inversion formula, cf. Proposition 2 [33]). *Let $f(x)$ be a real-valued function, defined for x ranging in a locally finite partially ordered set P . Let an element p exist with the property that $f(x) = 0$ unless $x \geq p$.*

Suppose that

$$g(x) = \sum_{y \leq x} f(y). \quad (1.4.1)$$

Then

$$f(x) = \sum_{y \leq x} g(y) \mu(y, x). \quad (1.4.2)$$

In the following proposition we have a similar formula as above for Möbius inversion formula.

Proposition 1.4.3 (General Möbius Inversion, cf. Proposition 2.6.3 of Peccati and Taqqu [31]). *Let S be a partially ordered set. For every pair of functions $G, F : S \rightarrow \mathbb{C}$ one has that, $\forall y \in S$,*

$$G(y) = \sum_{x \leq y} F(x) \quad \left(\text{resp. } G(y) = \sum_{y \leq x} F(x) \right) \quad (1.4.3)$$

if and only if, $\forall x \in S$,

$$F(x) = \sum_{y \preceq x} \mu(y, x)G(y) \quad \left(\text{resp. } F(x) = \sum_{x \preceq y} \mu(x, y)G(y) \right) \quad (1.4.4)$$

where μ is the Möbius function associated with S .

1.5 Information theory measures

Information theory is one of the prominent area of research which has many applications in science, engineering, finance, and economics. In this background on information theory part we cover basics of information theory. As a key reference we use "Elements of Information Theory" by Cover and Thomas [9] and "Information Theory and Network Coding" by Raymond W. Yeung [42]. We define entropy, joint entropy, conditional entropy, and mutual information for discrete random variables and their relations with each other. Then we carry our argument to continuous case by giving the definition of the differential entropy and we state the differential entropy of the multidimensional Gaussian distribution. Lastly, we give the definition of the probability entropy and then we state an existing result from literature related to the logarithmic Sobolev inequality.

Discrete case

Definition 1.5.1. Let X be a discrete random variable with probability mass function $p_X(x) = \mathbb{P}(X = x)$ then its Shannon entropy is defined by

$$H(X) = - \sum_x p_X(x) \log p_X(x) = -\mathbb{E}^{p_X(x)}[\log p_X(X)]. \quad (1.5.1)$$

Definition 1.5.2. Let (X, Y) be a pair of discrete random variables with a joint distribution $p(x, y) = \mathbb{P}(X = x, Y = y)$ then its joint entropy is defined by

$$H(X, Y) = - \sum_x \sum_y p(x, y) \log p(x, y) = -\mathbb{E}^{p(x, y)}[\log p(X, Y)] \quad (1.5.2)$$

and its conditional entropy is defined by

$$H(Y|X) = \sum_x p_X(x) H(Y|X = x) \quad (1.5.3)$$

$$\begin{aligned} &= - \sum_x p_X(x) \sum_y p(y|x) \log p(y|x) \\ &= - \sum_x \sum_y p(x, y) \log p(y|x) = -\mathbb{E}^{p(x, y)}[\log p(Y|X)]. \end{aligned} \quad (1.5.4)$$

Then we have the following relation between the joint and the conditional entropy:

$$H(X, Y) = H(X) + H(Y|X) = H(Y) + H(X|Y). \quad (1.5.5)$$

Definition 1.5.3. Let (X, Y) be a pair of discrete random variables with a joint distribution $p(x, y) = \mathbb{P}(X = x, Y = y)$ then its mutual information is defined by

$$\begin{aligned} I(X, Y) &= \sum_x \sum_y p(x, y) \log \frac{p(x, y)}{p_X(x)p_Y(y)} \\ &= \mathbb{E}^{p(x, y)} \left[\log \frac{p(X, Y)}{p_X(X)p_Y(Y)} \right]. \end{aligned}$$

We have following relations between mutual information and entropy:

$$I(X, Y) = H(X) + H(Y) - H(X, Y), \quad (1.5.6)$$

$$I(X, Y) = H(X) - H(X|Y) = H(Y) - H(Y|X), \quad (1.5.7)$$

$$I(X, Y) = I(Y, X), \quad (1.5.8)$$

$$I(X, X) = H(X). \quad (1.5.9)$$

Following two lemmas show that entropy and mutual information are submodular functions.

Recall that a function f is submodular iff $-f$ is supermodular.

Lemma 1.5.4 (Lemma 2,[27]). *Suppose $\{X_v\}_{v \in V}$ is a collection of random variables, and $f : 2^V \rightarrow \mathbf{R}$ is given by $f(S) = H(\{X_v\}_{v \in S})$, the entropy of the collection of random variables $\{X_v\}_{v \in V}$. Then f is a submodular function.*

Lemma 1.5.5 (Lemma 3,[27]). *Suppose $\{X_v\}_{v \in V}$ is a collection of random variables, and $f : 2^V \rightarrow \mathbf{R}$ is given by $f(A) = I(\{X_v\}_{v \in A}, \{X_v\}_{v \in V/A})$, the mutual information. Then f is a submodular function.*

Continuous case

Similarly to the discrete case we can define information theory measures for continuous random variables.

Definition 1.5.6. *Let X be a continuous random variable with probability density function $f_X(x)$ and support set S then its differential entropy is*

defined by

$$h(X) = - \int_S f_X(x) \log f_X(x) dx = -\mathbb{E}^{f_X(x)}[\log f_X(X)]. \quad (1.5.10)$$

Definition 1.5.7. Let (X, Y) be a pair of continuous random variables with a joint density $f(x, y)$ and support set S then its joint differential entropy is defined by

$$h(X, Y) = - \int_S f(x, y) \log f(x, y) dx dy \quad (1.5.11)$$

and its conditional differential entropy is defined by

$$h(X|Y) = - \int_S f(x, y) \log f(x|y) dx dy. \quad (1.5.12)$$

Then it is clear that

$$h(X, Y) = h(X|Y) + h(Y). \quad (1.5.13)$$

Definition 1.5.8. Let (X, Y) be a pair of continuous random variables with a joint density $f(x, y)$ and support set S then the mutual information is defined by

$$\begin{aligned} I(X, Y) &= \int_S f(x, y) \log \frac{f(x, y)}{f_X(x)f_Y(y)} dx dy \\ &= \mathbb{E}^{f(x, y)} \left[\log \frac{f(X, Y)}{f_X(X)f_Y(Y)} \right]. \end{aligned} \quad (1.5.14)$$

Then we have

$$I(X, Y) = h(X) - h(X|Y) = h(Y) - h(Y|X). \quad (1.5.15)$$

Entropy and differential entropy are nonnegative measures.

The differential entropy of the multivariate Gaussian distribution with mean $\vec{\mu}$ and covariance matrix Σ is given in the following proposition:

Proposition 1.5.9 (cf. Theorem 8.4.1 [9]). *Let $X \sim N(\vec{\mu}, \Sigma)$. Then we have*

$$h(X) = \frac{1}{2} \log((2\pi e)^d \det(\Sigma)).$$

1.6 Log-Sobolev inequality on the Poisson space

In this section, we use some notations of Zhang [43] to prepare ourselves for Section 6.5 where we derive log Sobolev inequalities for probability measures on the Poisson space. We also refer the reader to Ishikawa [15] for additional background on the stochastic calculus of variations for jump processes.

Before stating all necessary notations; we, first, want to state the definition of ‘the probability entropy’:

Definition 1.6.1. *The entropy of a random variable $F > 0$ is defined by*

$$\text{Ent}[F] = \mathbb{E}[F \log F] - \mathbb{E}[F] \log \mathbb{E}[F] \quad (1.6.1)$$

for sufficiently integrable F .

Let $\mathbb{U}$ be a Lusin space, that is, a Hausdorff space that is the image of a Polish space under a continuous bijection. We fix a σ -finite and infinite measure ν on $(\mathbb{U}, \mathcal{B}(\mathbb{U}))$.

Let Ω be the space of all non-negative integer-valued measures ω on $\mathbb{U} \times [0, 1]$ such that $\omega(\{u, t\}) \leq 1$ for any $(u, t) \in \mathbb{U} \times [0, 1]$ and $\omega(A \times [0, 1]) < +\infty$ for any $A \in \mathcal{B}(\mathbb{U})$ with $\nu(A) < \infty$. The canonical random measure μ_ω on Ω is then defined by

$$\mu_\omega(A \times (0, t]) = \omega(A \times (0, t]), \quad t \in [0, 1], \quad A \in \mathcal{B}(\mathbb{U}).$$

The filtration $(\mathcal{F}_t)_{t \in [0, 1]}$ is defined by

$$\mathcal{F}_t := \sigma\{\mu_\omega(A \times (0, s]) : s \leq t, A \in \mathcal{B}(\mathbb{U})\}.$$

We will simply denote $\mathcal{F}_1$ by $\mathcal{F}$. Let $\mathbb{P}$ be the probability measure on $(\Omega, \mathcal{F})$ such that μ_ω is a Poisson random measure with the intensity measure $\nu(du)$.

Denote by $\tilde{\mu}_\omega = \mu_\omega - \pi$ the compensated Poisson random measure, where $\pi(du, dt) := \nu(du) \times dt$ and dt is the Lebesgue measure on $[0, 1]$.

Let $\mathcal{F}_t^\mathbb{P}$ be the completion of $\mathcal{F}_t$ with respect to $\mathbb{P}$. Denote by $\mathcal{P}$ the predictable σ -field associated with $(\mathcal{F}_t^\mathbb{P})_{t \in [0, 1]}$, which is generated by all left continuous $\mathcal{F}_t^\mathbb{P}$ -adapted processes.

For $p \in [1, \infty]$ denote

$$\mathbb{L}^p := L^p \left\{ \mathbb{U} \times [0, 1] \times \Omega, \mathcal{B}(\mathbb{U} \times [0, 1]) \times \mathcal{F}^\mathbb{P}, \pi \times \mathbb{P} \right\}$$

and

$$\mathbb{L}_{\mathcal{P}}^p := L^p \{ \mathbb{U} \times [0, 1] \times \Omega, \mathcal{B}(\mathbb{U}) \times \mathcal{P}, \nu \times dt \times \mathbb{P} \}.$$

Let $\phi \in \mathbb{L}_{\mathcal{P}}^2$ be a bounded predictable process satisfying

$$\phi(u, t, \omega) \geq c_\phi > -1$$

and

$$\mathbb{E} \left[\left(\int_0^1 \int_{\mathbb{U}} \phi(u, t)^2 \pi(du, dt) \right)^2 \right] < \infty$$

and such that $t \rightarrow \mathcal{E}_t(\phi)$ is a square integrable $\mathcal{F}_t^{\mathbb{P}}$ -martingale.

The definition of $\mathcal{E}_t(\phi)$ is as follows

$$\begin{aligned} \mathcal{E}_t(\phi) := & \exp \left(\int_0^t \int_{\mathbb{U}} \log(1 + \phi(u, s)) \tilde{\mu}(du, ds) \right. \\ & \left. + \int_0^t \int_{\mathbb{U}} (\log(1 + \phi(u, s)) - \phi(u, s)) \pi(du, ds) \right). \end{aligned} \quad (1.6.2)$$

The $\mathcal{E}_t(\phi)$ is well defined due to the inequalities

$$|\log(1 + x)| \leq C|x| \text{ and } |\log(1 + x) - x| \leq C|x|^2 \text{ for some } C > 0.$$

We also let

$$\|\phi\|_2^2 = \int_0^1 \int_{\mathbb{U}} |\phi(u, s)|^2 \pi(du, ds).$$

Let D denote the finite difference operator defined by

$$D_{u,s}F(\omega) = F(\omega \cup \{(u, s)\}) - F(\omega),$$

$(u, s) \in \mathbb{U} \times [0, 1]$, $\omega \in \Omega$, for any random variable $F : \Omega \rightarrow \mathbb{R}$.

Bobkov and Götze [7] showed logarithmic Sobolev inequality in the Poisson space for the first time: Let π_θ be the Poisson measure on $\mathbf{N}$ with parameter $\theta > 0$. Then

$$\text{Ent}_{\pi_\theta}(F) \leq \theta \cdot \mathbb{E}^{\pi_\theta} \left[\frac{|DF|^2}{F} \right] \quad (1.6.3)$$

for any $F : \mathbf{N} \rightarrow (0, \infty)$; and

$$\text{Ent}_{\pi_\theta}(e^F) \leq \theta \cdot \mathbb{E}^{\pi_\theta} [e^F (|DF| e^{|DF|} - e^{|DF|} + 1)] \quad (1.6.4)$$

for any $F : \mathbf{N} \rightarrow \mathbf{R}$ bounded. If D were a differentiable operator (1.6.3) would be equivalent to the usual L^1 -logarithmic Sobolev inequality:

$$\text{Ent}_{\pi_\theta}(F) \leq \theta \cdot \mathbb{E}^{\pi_\theta} [DF \cdot D \log F] \quad (1.6.5)$$

for any $F : \mathbf{N} \rightarrow (0, \infty)$. Therefore (1.6.3) and (1.6.4) are called *modified logarithmic Sobolev inequalities*.

By martingale representation approach, Wu [41] established the following modified logarithmic Sobolev inequality: for any $F : \mathbf{N} \rightarrow (0, \infty)$

$$\text{Ent}_{\pi_\theta}(F) \leq \mathbb{E}^{\pi_\theta} [D\Phi(F) - \Phi'(F)DF] \quad (1.6.6)$$

where $\Phi(x) = x \log x$ covers (1.6.3), (1.6.4), and (1.6.5) as particular cases.

He showed in Theorem 1.1 of [41] for any $F \in L^1(W, \mathcal{P})$ s.t. $F > 0$

$$\begin{aligned} \text{Ent}_{\mathcal{P}}(F) &\leq \mathbb{E}^{\mathcal{P}} \left[\int_E [D_z \Phi(F) - \Phi'(F) D_z F] \nu(dz) \right] \\ &= \mathbb{E}^{\mathcal{P}} \left[\int_E [(F + D_z F) \log(F + D_z F) - (F + D_z F) \log(F) - D_z F] \nu(dz) \right]. \end{aligned}$$

(1.6.7)

We see (1.6.7) contains (1.6.6) as a particular case, so (1.6.7) contains (1.6.3), (1.6.4), and (1.6.5) as particular cases.

Part I

Stochastic and supermodular ordering

Chapter 2

Supermodular ordering of Poisson arrays

Summary of the Chapter

We derive necessary and sufficient conditions for the supermodular ordering of certain triangular arrays of Poisson random variables, based on the componentwise ordering of their covariance matrices. Applications are proposed for markets driven by jump-diffusion processes, using sums of Gaussian and Poisson random vectors. In addition, applications will be proposed for markets driven by Ornstein-Uhlenbeck and jump-diffusion processes in Chapter 4. Our results rely on a new triangular structure for the representation of Poisson random vectors using their Lévy-Khintchine representation.

2.1 Introduction

Stochastic ordering of random vectors is used in finance and economics as a risk management tool that yields finer information than mean-variance analysis cf. [21] and references therein for a review with economic interpretations.

It is well-known (cf. e.g. Theorem 3.9.5 of [26])(see Theorem 1.1.6) that necessary conditions for $X \leq_{\text{sm}} Y$ are:

- (i) X_i and Y_i have same distribution for all $i = 1, \dots, d$, and
- (ii) for all $1 \leq i < j \leq d$ we have

$$\text{Cov}(X_i, X_j) \leq \text{Cov}(Y_i, Y_j),$$

where (i) above follows from the fact that any function of a single variable is supermodular, and (ii) follows by application of (1.1.1) to the supermodular function $\Phi_{i,j}(x_1, \dots, x_d) := (x_i - \mathbb{E}[X_i])(x_j - \mathbb{E}[X_j])$, $1 \leq i < j \leq d$. The definitions of supermodular ordering and supermodular functions are given in Definition 1.1.2.

Supermodular ordering of Gaussian random vectors has been completely characterized in [25] Theorem 4.2 (see Theorem 1.1.7), cf. also Theorem 3.13.5 of [26] (see Theorem 1.1.7), by showing that (i) and (ii) above are also sufficient conditions when X and Y are Gaussian, cf. also Müller [24] for other orderings (stochastic, convex, convex increasing, supermodular) of Gaussian random vectors.

In this chapter we consider the supermodular ordering of vectors of Poisson random variables, see Teicher [40] and [16] for early references on the multivariate Poisson distribution. As noted in Section 2.2, any d -dimensional Poisson random vector is based on $2^d - 1$ parameters, therefore (i) and (ii), which are based on $d(d+1)/2$ conditions, cannot characterize their distribution ordering except if $d = 2$.

For this reason, in Section 2.3 we consider a particular dependence structure of Poisson arrays depending on $d(d+1)/2$ parameters based on a natural decomposition of their Lévy measure on the vertices of the d -dimensional unit hypercube. We show in Theorem 2.4.2 that in this case, conditions (i) and (ii) become necessary and sufficient for the supermodular ordering of X and Y as in the Gaussian setting. When $d = 2$, no restriction has to be imposed on 2-dimensional Poisson random vectors X and Y . Triangular Poisson structures of a different type have been considered in Sim [37] for the simulation of Poisson random vectors.

Our proof relies on the characterization of the supermodular ordering of d -dimensional infinitely divisible random vectors X and Y by their Lévy measures, cf. [3], based on the covariance identities of Houdré et al. [14]. Extensions of such identities have already been applied to other multidimensional stochastic (including convex) orderings in e.g. [5] based on stochastic calculus and in Arnaudon et al. [2] using forward-backward stochastic calculus.

We also derive sufficient conditions for the supermodular ordering of sums of Gaussian and Poisson random vectors, with application to a jump-diffusion asset price model, cf. Theorem 2.4.4. Indeed, the supermodular ordering of random asset vectors implies the stop-loss ordering of their associated portfolios, cf. Theorem 3.1 in [23] or Theorem 8.3.3 in [26] (see Theorem 1.1.8).

We proceed as follows. In Section 2.2 we recall the construction of Poisson random vectors, and in Section 2.3 we specialize this construction to a certain dependence setting based on Poisson arrays. Finally in Section 2.4 we prove our main characterization of supermodular ordering for such vectors, including extensions to the increasing supermodular order, cf. Theorem 2.4.2 and Proposition 2.4.3. In addition we provide a sufficient condition for the supermodular ordering of sums of Gaussian and Poisson random vectors in Theorem 2.4.4, with application to an exponential jump-diffusion asset price model. We also include a remark on the related convex ordering problem for such Poisson arrays in Proposition 2.4.5.

2.2 Poisson random vectors

Consider $X = (X_i)_{1 \leq i \leq d}$ a d -dimensional infinitely divisible Poisson random vector with Lévy measure μ on $\mathbb{R}^d$, which satisfies

$$E[e^{i\langle \bar{t}, X \rangle}] = \exp \left(\int_{\mathbb{R}^d} (e^{i\langle \bar{t}, x \rangle} - 1) \mu(dx) \right),$$

where $\bar{t} = (t_1, \dots, t_d) \in \mathbb{R}^d$ and $\langle \cdot, \cdot \rangle$ denotes the scalar product in $\mathbb{R}^d$.

Since $(X_i)_{1 \leq i \leq d}$ has Poisson marginals, all marginals of μ on $\mathbb{R}^d$ are supported by $\{0, 1\}$, and consequently the Lévy measure $\mu(dx)$ is supported on the vertices of the unit hypercube of $\mathbb{R}^d$ and takes the form

$$\mu(dx) = \sum_{\emptyset \neq S \subset \{1, 2, \dots, d\}} a_S \delta_{e_S}(dx),$$

where $(a_S)_{\emptyset \neq S \subset \{1, 2, \dots, d\}}$ is a family of nonnegative numbers and

$$C_d = \{0, 1\}^d = \left\{ e_S := \sum_{i \in S} e_i : S \subset \{1, \dots, d\} \right\}$$

denotes the vertices of the d -dimensional unit hypercube, identified to the power set $\{0, 1\}^d \simeq \{S \subset \{1, \dots, d\}\}$ of $\{1, \dots, d\}$, and $(e_k)_{1 \leq k \leq d}$ is the canonical basis of $\mathbb{R}^d$.

Consequently, any d -dimensional Poisson random vector $X = (X_1, \dots, X_d)$ can be represented as

$$X_i = \sum_{\substack{S \subset \{0, 1\}^d \\ S \neq \emptyset}} \mathbf{1}_{\{i \in S\}} X_S = \sum_{\substack{S \subset \{1, 2, \dots, d\} \\ S \ni i}} X_S, \quad i = 1, \dots, d, \quad (2.2.1)$$

where $(X_S)_{\emptyset \neq S \subset \{1, 2, \dots, d\}}$ is a family of $2^d - 1$ independent Poisson random variables with respective intensities $(a_S)_{\emptyset \neq S \subset \{1, 2, \dots, d\}}$, cf. also Theorem 3 of [16] (see Theorem 1.2.1).

In particular, at most $2^d - 1$ degrees of freedom are needed in order

to characterize the probability distribution of X . Note also that

$$\text{Cov}(X_i, X_j) = \sum_{\substack{S \subset \{1,2,\dots,d\} \\ S \ni i,j}} \text{Var}[X_S] \geq 0, \quad i, j = 1, \dots, d,$$

cf. also Theorem 4 of [16] (see Theorem 1.2.2), showing that X has positive covariances.

For example when $d = 2$ we have

$$\begin{cases} X_1 = X_{\{1\}} + X_{\{1,2\}} \\ X_2 = X_{\{2\}} + X_{\{1,2\}}, \end{cases} \quad (2.2.2)$$

and when $d = 3$ we find

$$\begin{cases} X_1 = X_{\{1\}} + X_{\{1,2\}} + X_{\{1,3\}} + X_{\{1,2,3\}} \\ X_2 = X_{\{2\}} + X_{\{1,2\}} + X_{\{2,3\}} + X_{\{1,2,3\}} \\ X_3 = X_{\{3\}} + X_{\{1,3\}} + X_{\{2,3\}} + X_{\{1,2,3\}}. \end{cases} \quad (2.2.3)$$

By decomposing the sum (2.2.1) according to the $\binom{d}{k}$ subsets $S_i^1, \dots, S_i^{\binom{d}{k}}$ of $\{1, \dots, d\}$ of sizes $k = 1, \dots, d$ we can also write

$$X_i = \sum_{k=1}^d \sum_{j=1}^{\binom{d}{k}} \mathbf{1}_{\{i \in S_k^j\}} X_{S_k^j}, \quad i = 1, \dots, d.$$

2.3 Triangular Poisson arrays

Since Poisson random vectors can depend on $2^d - 1$ degrees of freedom $(a_S)_{\emptyset \neq S \subset \{1,2,\dots,d\}}$, it should be generally not possible to characterize their

CHAPTER 2. SUPERMODULAR ORDERING OF POISSON ARRAYS

ordering based on the data of its covariance matrix which contains only $d(d+1)/2$ components.

In this chapter we restrict ourselves to d -dimensional Poisson random vectors $(X_i)_{1 \leq i \leq d}$ whose Lévy measure $\mu(dx)$ is supported by

$$\{\mathbf{e}_{i,j} := e_1 + \cdots + e_{i-1} + e_j : 1 \leq i \leq j \leq d\},$$

i.e. $\mu(dx)$ takes the form

$$\mu(dx) = \sum_{1 \leq i \leq j \leq d} a_{i,j} \delta_{\mathbf{e}_{i,j}}(dx), \quad (2.3.1)$$

where

$$\mathbf{e}_{i,j} = (1, \dots, \underset{\substack{\uparrow \\ i-1}}{1}, 0, \dots, 0, \underset{\substack{\uparrow \\ j}}{1}, 0, \dots, 0), \quad 1 \leq i \leq j \leq d,$$

and the vector $(X_i)_{1 \leq i \leq d}$ actually depends on $d(d+1)/2$ parameters $(a_{i,j})_{1 \leq i \leq j \leq d}$.

In other words the d -dimensional Poisson random vector $(X_i)_{1 \leq i \leq d}$ with Lévy measure μ given by (2.3.1) can be represented as

$$X = \sum_{1 \leq k \leq l \leq d} X_{k,l} \mathbf{e}_{k,l}, \quad (2.3.2)$$

where $(X_{i,j})_{1 \leq i \leq j \leq d}$ is a triangular array of independent Poisson random

variables with respective parameters $(a_{i,j})_{1 \leq i \leq j \leq d}$, and we have

$$X_i = \sum_{k=1}^i X_{k,i} + \sum_{k=i+1}^d \sum_{l=k}^d X_{k,l}, \quad (2.3.3)$$

$i = 1, \dots, d$.

For example when $d = 2$ we have

$$\begin{cases} X_1 = X_{1,1} + X_{2,2} \\ X_2 = X_{1,2} + X_{2,2}, \end{cases} \quad (2.3.4)$$

which coincides with (2.2.2), and when $d = 3$ we get

$$\begin{cases} X_1 = X_{1,1} + X_{2,2} + X_{2,3} + X_{3,3} \\ X_2 = X_{1,2} + X_{2,2} + X_{3,3} \\ X_3 = X_{1,3} + X_{2,3} + X_{3,3}, \end{cases}$$

which, in comparison with (2.2.3), excludes an interaction component specific to rows 2 and 3. Note that the 2-dimensional case (2.3.4) does not impose any such restriction on the form of the vector X .

From (2.3.3) we have

$$\text{Cov}(X_i, X_i) = \text{Var}(X_i) = \mathbb{E}[X_i] = \sum_{k=1}^i a_{k,i} + \sum_{k=i+1}^d \sum_{l=k}^d a_{k,l}, \quad (2.3.5)$$

and

$$\text{Cov}(X_i, X_j) = \sum_{k=i+1}^j a_{k,j} + \sum_{k=j+1}^d \sum_{l=k}^d a_{k,l}, \quad 1 \leq i < j \leq d, \quad (2.3.6)$$

hence the vector $C = (\text{Cov}(X_i, X_j))_{1 \leq i \leq j \leq d}$ is obtained from $A = (a_{k,l})_{1 \leq k \leq l \leq d}$ and the matrix transformation $C = M \times A$, i.e.

$$\text{Cov}(X_i, X_j) = \sum_{1 \leq k \leq l \leq d} a_{k,l} M_{i,j,k,l}, \quad 1 \leq i \leq j \leq d,$$

where the $(d(d+1)/2)^2$ -dimensional matrix $M = (M_{i,j,k,l})_{\substack{1 \leq i \leq j \leq d \\ 1 \leq k \leq l \leq d}}$ is given by

$$M_{i,j,k,l} = \mathbf{1}_{\{j+1 \leq k \leq l \leq d\}} + \mathbf{1}_{\{1 \leq k \leq i, i=j=l\}} + \mathbf{1}_{\{i+1 \leq k \leq j, i < j=l\}}.$$

We can check by an elementary computation that the inverse M^{-1} of M is given by the inversion formula

$$a_{k,l} = \begin{cases} \text{Cov}(X_1, X_1) + \sum_{j=3}^d \text{Cov}(X_2, X_j) - \sum_{j=2}^d \text{Cov}(X_1, X_j), & k = l = 1, \\ \text{Cov}(X_l, X_l) - \text{Cov}(X_1, X_l), & k = 1, 2 \leq l \leq d, \\ \text{Cov}(X_{k-1}, X_l) - \text{Cov}(X_k, X_l), & 2 \leq k < l \leq d, \\ \text{Cov}(X_{k-1}, X_k) + \sum_{j=k+2}^d \text{Cov}(X_{k+1}, X_j) - \sum_{j=k+1}^d \text{Cov}(X_k, X_j), & 2 \leq k = l \leq d, \end{cases} \quad (2.3.7)$$

with, for $d = 2$,

$$\begin{cases} a_{1,1} = \text{Cov}(X_1, X_1) - \text{Cov}(X_1, X_2), \\ a_{1,2} = \text{Cov}(X_2, X_2) - \text{Cov}(X_1, X_2), \\ a_{2,2} = \text{Cov}(X_1, X_2). \end{cases}$$

In particular it follows from (2.3.7) that the distribution of X given by (2.3.3) is characterized by its covariance matrix $(\text{Cov}(X_i, X_j))_{1 \leq i \leq j \leq d}$, and the vector X admits the representation (2.3.2) provided its covariances yield non-negative values of $(a_{k,l})_{1 \leq k \leq l \leq d}$ in (2.3.7). Such conditions differ from the ones imposed in relation with the triangular representation of [37].

In the simple case $a_{i,j} = 0$, $1 \leq i < j \leq d$, we have

$$X_i = \sum_{k=i}^d X_{k,k}, \quad i = 1, \dots, d,$$

and $a_{i,i} = \text{Var}[X_i] = \text{Var}[X_{i,i}] + \dots + \text{Var}[X_{d,d}]$, $i = 1, \dots, d$.

2.4 Supermodular Poisson arrays

When $d = 2$ we note that from (2.3.3)-(2.3.6) we have

$$\mathbb{E}[X_1] = a_{1,1} + a_{2,2}, \quad \mathbb{E}[X_2] = a_{1,2} + a_{2,2}, \quad \text{and} \quad \text{Cov}(X_1, X_2) = a_{2,2},$$

hence the Lévy measure $\mu(dx)$ of any 2-dimensional Poisson random vector $X = (X_1, X_2)$ can be represented as

$$\begin{aligned} \int_{\mathbb{R}^2} \phi(x) \mu(dx) &= \phi(1, 0)a_{1,1} + \phi(0, 1)a_{1,2} + \phi(1, 1)a_{2,2} \\ &= \phi(1, 0)(\mathbb{E}[X_1] - \text{Cov}(X_1, X_2)) + \phi(0, 1)(\mathbb{E}[X_2] - \text{Cov}(X_1, X_2)) + \phi(1, 1)\text{Cov}(X_1, X_2) \\ &= E[X_1]\phi(1, 0) + E[X_2]\phi(0, 1) + \text{Cov}(X_1, X_2)(\phi(1, 1) - \phi(1, 0) - \phi(0, 1)), \end{aligned}$$

for any function $\phi : C_2 \longrightarrow \mathbb{R}$ such that $\phi(0) = 0$.

More generally, for all $d \geq 2$, the next Lemma 2.4.1 provides a representation of Lévy measures which will be used in the proof of our main result Theorem 2.4.2 below.

Lemma 2.4.1. *Assume that X is a Poisson array represented as in (2.3.2), with Lévy measure μ . Then for any function $\phi : C_d \longrightarrow \mathbb{R}$ such that $\phi(0) = 0$ we have*

$$\begin{aligned} \int_{\mathbb{R}^d} \phi(x) \mu(dx) &= \sum_{j=1}^d E[X_j] \phi(\mathbf{e}_{1,j}) \\ &\quad + \sum_{1 \leq i < j \leq d} \text{Cov}(X_i, X_j) (\phi(\mathbf{e}_{i+1,j}) + \phi(\mathbf{e}_{i-1,i-1}) - \phi(\mathbf{e}_{i,i}) - \phi(\mathbf{e}_{i,j})), \end{aligned}$$

with $\mathbf{e}_{0,0} := (0, \dots, 0)$.

Proof. We have

$$\begin{aligned} &\sum_{1 \leq i < j \leq d} \text{Cov}(X_i, X_j) (\phi(\mathbf{e}_{i+1,j}) + \phi(\mathbf{e}_{i-1,i-1}) - \phi(\mathbf{e}_{i,i}) - \phi(\mathbf{e}_{i,j})) \\ &= \left(\sum_{j=3}^d \text{Cov}(X_2, X_j) - \sum_{j=2}^d \text{Cov}(X_1, X_j) \right) \phi(\mathbf{e}_{1,1}) - \sum_{j=2}^d \text{Cov}(X_1, X_j) \phi(\mathbf{e}_{1,j}) \end{aligned}$$

$$\begin{aligned}
& + \sum_{i=2}^d \left(\text{Cov}(X_{i-1}, X_i) + \sum_{j=i+2}^d \text{Cov}(X_{i+1}, X_j) - \sum_{j=i+1}^d \text{Cov}(X_i, X_j) \right) \phi(\mathbf{e}_{i,i}) \\
& + \sum_{2 \leq i < j \leq d} (\text{Cov}(X_{i-1}, X_j) - \text{Cov}(X_i, X_j)) \phi(\mathbf{e}_{i,j}) \\
& = - \sum_{j=1}^d E[X_j] \phi(\mathbf{e}_{1,j}) + \sum_{1 \leq i \leq j \leq d} a_{i,j} \phi(\mathbf{e}_{i,j}),
\end{aligned}$$

where the conclusion comes from the above inversion formula (2.3.7). $\square$

The above Lemma 2.4.1 shows that the transpose $M^{-1\dagger}$ of M^{-1} is given by

$$M_{i,j,k,l}^{-1\dagger} = \mathbf{1}_{\{k=1 \leq i=j=l\}} + \mathbf{1}_{\{k=i+1 \leq l=j\}} + \mathbf{1}_{\{k=l=i-1 < j-1\}} - \mathbf{1}_{\{k=l=i < j\}} - \mathbf{1}_{\{k=i < l=j\}},$$

and we have

$$\begin{aligned}
\int_{\mathbb{R}^d} \phi(x) \mu(dx) &= \sum_{1 \leq k \leq l \leq d} a_{k,l} \phi(\mathbf{e}_{k,l}) \\
&= \sum_{1 \leq k \leq l \leq d} \phi(\mathbf{e}_{k,l}) \sum_{1 \leq i \leq j \leq d} M_{k,l,i,j}^{-1} \text{Cov}(X_i, X_j) \\
&= \sum_{1 \leq i \leq j \leq d} \text{Cov}(X_i, X_j) \sum_{1 \leq k \leq l \leq d} M_{k,l,i,j}^{-1} \phi(\mathbf{e}_{k,l}) \\
&= \sum_{1 \leq i \leq j \leq d} \text{Cov}(X_i, X_j) \sum_{1 \leq k \leq l \leq d} M_{i,j,k,l}^{-1\dagger} \phi(\mathbf{e}_{k,l}).
\end{aligned}$$

Consider now two Poisson arrays X and Y whose respective Lévy measures μ and ν are represented as

$$\mu(dx) = \sum_{1 \leq i \leq j \leq d} a_{i,j} \delta_{\mathbf{e}_{i,j}}(dx) \quad \text{and} \quad \nu(dx) = \sum_{1 \leq i \leq j \leq d} b_{i,j} \delta_{\mathbf{e}_{i,j}}(dx),$$

as in (2.3.1). If X_i has the same distribution as Y_i for all $i = 1, \dots, d$

then $E[X_i] = E[Y_i]$, $i = 1, \dots, d$, and Lemma 2.4.1 shows that

$$\begin{aligned}
 & \int_{\mathbb{R}^d} \phi(x) \nu(dx) - \int_{\mathbb{R}^d} \phi(x) \mu(dx) \\
 &= \sum_{1 \leq i < j \leq d} (\text{Cov}(Y_i, Y_j) - \text{Cov}(X_i, X_j)) (\phi(\mathbf{e}_{i+1,j}) + \phi(\mathbf{e}_{i-1,i-1}) - \phi(\mathbf{e}_{i,i}) - \phi(\mathbf{e}_{i,j})) \\
 &= \sum_{1 \leq i < j \leq d} (\text{Cov}(Y_i, Y_j) - \text{Cov}(X_i, X_j)) t_{i,j} \phi(\mathbf{e}_{i-1,i-1}), \tag{2.4.1}
 \end{aligned}$$

where $t_{i,j}$ is the linear mapping defined on ϕ by

$$t_{i,j} \phi(z) = \phi(z + e_i + e_j) + \phi(z) - \phi(z + e_i) - \phi(z + e_j),$$

and

$$\begin{aligned}
 t_{i,j} \phi(\mathbf{e}_{i-1,i-1}) &= \phi(\mathbf{e}_{i-1,i-1} + e_i + e_j) + \phi(\mathbf{e}_{i-1,i-1}) - \phi(\mathbf{e}_{i-1,i-1} + e_i) - \phi(\mathbf{e}_{i-1,i-1} + e_j) \\
 &= \phi(\mathbf{e}_{i+1,j}) + \phi(\mathbf{e}_{i-1,i-1}) - \phi(\mathbf{e}_{i,i}) - \phi(\mathbf{e}_{i,j}).
 \end{aligned}$$

In other words, (2.4.1) provides the explicit form of the decomposition stated in Theorem 1 of [21] (see Theorem 1.1.9).

Relation (2.4.1) also shows that the nonnegativity of the coefficients

$$\text{Cov}(Y_i, Y_j) - \text{Cov}(X_i, X_j) \geq 0, \quad 1 \leq i < j \leq d, \tag{2.4.2}$$

becomes a necessary and sufficient condition for the supermodular ordering of the Lévy measures μ and ν which here have *finite* support.

Next is the main result of this chapter. It reformulates (2.4.2) as a

necessary and sufficient condition for supermodular ordering of triangular Poisson arrays, based on Theorem 4.5 of [3] (see Theorem 1.3.3) which allows us to carry over the notion of supermodularity from the finite support setting of Lévy measures μ, ν on the cube C_d , to the infinite support setting of Poisson random variables.

Theorem 2.4.2. *Consider two Poisson arrays X and Y both represented as in (2.3.2). Then the conditions*

$$\mathbb{E}[X_i] = \mathbb{E}[Y_i], \quad 1 \leq i \leq d, \quad (2.4.3)$$

and

$$\text{Cov}(X_i, X_j) \leq \text{Cov}(Y_i, Y_j), \quad 1 \leq i < j \leq d, \quad (2.4.4)$$

are necessary and sufficient for the supermodular ordering $X \leq_{\text{sm}} Y$.

Proof. By Theorem 4.5 in [3] (see Theorem 1.3.3) it suffices to show that

$$\int_{\mathbb{R}^d} \phi(x) \mu(dx) \leq \int_{\mathbb{R}^d} \phi(y) \nu(dy)$$

for all supermodular functions $\phi : \mathbb{R}^d \rightarrow \mathbb{R}$ satisfying $\phi(0) = \phi(\mathbf{e}_{0,0}) = 0$, where $\mu(dx)$ and $\nu(dy)$ respectively denote the Lévy measures of X and Y . By Lemma 2.4.1 we have the identity

$$\begin{aligned} & \int_{\mathbb{R}^d} \phi(x) \nu(dx) - \int_{\mathbb{R}^d} \phi(x) \mu(dx) \\ &= \sum_{1 \leq i < j \leq d} (\text{Cov}(Y_i, Y_j) - \text{Cov}(X_i, X_j)) (\phi(\mathbf{e}_{i+1,j}) + \phi(\mathbf{e}_{i-1,i-1}) - \phi(\mathbf{e}_{i,i}) - \phi(\mathbf{e}_{i,j})). \end{aligned} \quad (2.4.5)$$

CHAPTER 2. SUPERMODULAR ORDERING OF POISSON ARRAYS

Next, the supermodular inequalities

$$\phi(\mathbf{e}_{i,i}) + \phi(\mathbf{e}_{i,j}) \leq \phi(\mathbf{e}_{i,i} \wedge \mathbf{e}_{i,j}) + \phi(\mathbf{e}_{i,i} \vee \mathbf{e}_{i,j}) = \phi(\mathbf{e}_{i-1,i-1}) + \phi(\mathbf{e}_{i+1,j}) \quad (2.4.6)$$

show that

$$(\text{Cov}(Y_i, Y_j) - \text{Cov}(X_i, X_j))(\phi(\mathbf{e}_{i+1,j}) + \phi(\mathbf{e}_{i-1,i-1}) - \phi(\mathbf{e}_{i,i}) - \phi(\mathbf{e}_{i,j})) \geq 0, \quad 1 \leq i < j \leq d,$$

under condition (2.4.4), which allows us to conclude by (2.4.5). $\square$

Note that in view of Condition (2.4.3), the inequality (2.4.4) can be rewritten using joint moments, as

$$E[X_i X_j] \leq E[Y_i Y_j], \quad 1 \leq i < j \leq d.$$

The next proposition is relative to the increasing supermodular order.

Proposition 2.4.3. *Consider two Poisson arrays X and Y both represented as in (2.3.2), and assume that*

$$\mathbb{E}[X_i] \leq \mathbb{E}[Y_i], \quad 1 \leq i \leq d,$$

and

$$\text{Cov}(X_i, X_j) \leq \text{Cov}(Y_i, Y_j), \quad 1 \leq i < j \leq d.$$

Then we have

$$\mathbb{E}[\Phi(X)] \leq \mathbb{E}[\Phi(Y)]$$

for all nondecreasing supermodular functions $\Phi : \mathbb{R}^d \rightarrow \mathbb{R}$.

Proof. First we note that Theorem 4.6 of [3] (see Theorem 1.3.3) can be easily extended to nondecreasing supermodular functions ϕ on $\mathbb{R}^d$ satisfying $\phi(0) = 0$, using the same approximation as in Lemma 4.3 therein. The conclusion follows again from Lemma 2.4.1 and the decomposition

$$\begin{aligned} & \int_{\mathbb{R}^d} \phi(x) \nu(dx) - \int_{\mathbb{R}^d} \phi(x) \mu(dx) \\ &= \sum_{j=1}^d (E[Y_j] - E[X_j]) \phi(\mathbf{e}_{1,j}) \\ & \quad + \sum_{1 \leq i < j \leq d} (\text{Cov}(Y_i, Y_j) - \text{Cov}(X_i, X_j)) (\phi(\mathbf{e}_{i+1,j}) + \phi(\mathbf{e}_{i-1,i-1}) - \phi(\mathbf{e}_{i,i}) - \phi(\mathbf{e}_{i,j})), \end{aligned}$$

for all *nondecreasing* supermodular functions $\phi : C_d \rightarrow \mathbb{R}$, since $\phi(\mathbf{e}_{1,j}) \geq \phi(0) = 0$, $j = 1, \dots, d$. $\square$

Sums of Gaussian and Poisson vectors

Theorem 2.4.2 can be extended to a sufficient condition for the sum of Gaussian and Poisson vectors, as follows.

Theorem 2.4.4. *Consider two Poisson arrays X and Y both represented as in (2.3.2), and two d -dimensional centered Gaussian random vectors U, V such that U is independent of X and V is independent of Y . Then the conditions $\mathbb{E}[X_i] = \mathbb{E}[Y_i]$, $1 \leq i \leq d$, and*

$$\text{Cov}(U_i, U_j) \leq \text{Cov}(V_i, V_j), \quad \text{Cov}(X_i, X_j) \leq \text{Cov}(Y_i, Y_j),$$

are sufficient for the supermodular ordering $X + U \leq_{\text{sm}} Y + V$.

Proof. This is a consequence of Theorem 4.2 of [25] (see Theorem 1.1.7)

CHAPTER 2. SUPERMODULAR ORDERING OF POISSON ARRAYS

on Gaussian random vectors, Theorem 2.4.2 above, and the fact that the supermodular ordering is closed under convolution, cf. Theorem 3.9.14-(C) of [26] (see Theorem 1.1.4). $\square$

Proposition 2.4.3 can be similarly extended to sums of Gaussian vectors and Poisson arrays.

As an application of Theorem 2.4.4 we can consider exponential jump-diffusion models $(X_t^1, \dots, X_t^d)_{t \in \mathbb{R}_+}$ of the form

$$X_t^i = X_0^i \exp(\sigma_i B_t^i + \eta_i Z_t^i + \mu_i t), \quad i = 1, \dots, d, \quad (2.4.7)$$

as in (2.3.3), where Z_t^i is the Poisson jump process

$$Z_t^i = \sum_{k=1}^i N_t^{k,i} + \sum_{k=i+1}^d \sum_{l=k}^d N_t^{k,l}, \quad i = 1, \dots, d,$$

and $\{(N_t^{k,l})_{t \in \mathbb{R}_+}, 1 \leq k \leq l \leq d\}$ is a family of $d(d+1)/2$ independent standard Poisson processes with respective intensities $(a_{k,l})_{1 \leq k \leq l \leq d}$, $(B_t^1, \dots, B_t^d)_{t \in \mathbb{R}_+}$ is a d -dimensional Brownian motion, X_0^i, σ_i, η_i are non-negative coefficients, and $\mu_i \in \mathbb{R}, i = 1, \dots, d$.

In that case, Theorem 2.4.4 shows that the supermodular ordering of vectors of the form $(X_t^1, \dots, X_t^d)$ follows from the componentwise orderings of the covariance matrices of the d -dimensional Brownian and Poisson processes $(B_t^1, \dots, B_t^d)$ and $(Z_t^1, \dots, Z_t^d)$, cf. Theorem 3.9.8-(P8) of [26] (see Theorem 1.1.5), since every exponential function in (2.4.7) is increasing.

Convex ordering

We end this chapter by some comments in the case of the convex ordering.

Proposition 2.4.5. *Consider two Poisson arrays X and Y both represented as in (2.3.2). Then we have $X \leq_{\text{cx}} Y$ if and only if X and Y have same distributions.*

Proof. Assume $X \leq_{\text{cx}} Y$, i.e. we have

$$\mathbb{E}[\Phi(X)] \leq \mathbb{E}[\Phi(Y)]$$

for all *convex* functions $\Phi : \mathbb{R}^d \rightarrow \mathbb{R}$. Clearly this implies $E[X_k] = E[Y_k]$, $k = 1, \dots, d$. Next, letting $0 \leq k < l \leq d$ and assuming that $\text{Cov}(Y_i, Y_j) - \text{Cov}(X_i, X_j) = 0$ whenever $1 \leq i \leq k \leq d$ and $1 \leq i < j \leq l \leq d$, we check that the function

$$(x_1, \dots, x_d) \mapsto \phi_{k,l+1}(x_1, \dots, x_d) := \max \left(0, 2x_{l+1} - \sum_{a=k}^d x_a \right)$$

is convex on $\mathbb{R}^d$. In addition, taking again $\mathbf{e}_{i,j} := e_1 + \dots + e_{i-1} + e_j$, $1 \leq i \leq j \leq d$, we note that it satisfies $\phi_{k,l+1}(\mathbf{e}_{i,j}) = 1$ if and only if $1 \leq i \leq k < l+1 = j \leq d$, otherwise $\phi_{k,l+1}(\mathbf{e}_{i,j}) = 0$, hence

$$\phi_{k,l+1}(\mathbf{e}_{k+1,l+1}) + \phi_{k,l+1}(\mathbf{e}_{k-1,k-1}) - \phi_{k,l+1}(\mathbf{e}_{k,k}) - \phi_{k,l+1}(\mathbf{e}_{k,l+1}) < 0,$$

while

$$\phi_{k,l+1}(\mathbf{e}_{i+1,j}) + \phi_{k,l+1}(\mathbf{e}_{i-1,i-1}) - \phi_{k,l+1}(\mathbf{e}_{i,i}) - \phi_{k,l+1}(\mathbf{e}_{i,j}) = 0$$

CHAPTER 2. SUPERMODULAR ORDERING OF POISSON ARRAYS

whenever $1 \leq k < i \leq d$ or $2 \leq l+1 < j \leq d$. Hence from Lemma 2.4.1, the condition $\text{Cov}(Y_k, Y_{l+1}) > \text{Cov}(X_k, X_{l+1})$ implies

$$\begin{aligned}
& \int_{\mathbb{R}^d} \phi(x) \nu(dx) - \int_{\mathbb{R}^d} \phi(x) \mu(dx) \\
&= \sum_{1 \leq i < j \leq d} (\text{Cov}(Y_i, Y_j) - \text{Cov}(X_i, X_j)) (\phi(\mathbf{e}_{i+1,j}) + \phi(\mathbf{e}_{i-1,i-1}) - \phi(\mathbf{e}_{i,i}) - \phi(\mathbf{e}_{i,j})) \\
&= (\text{Cov}(Y_k, Y_{l+1}) - \text{Cov}(X_k, X_{l+1})) (\phi(\mathbf{e}_{k+1,l+1}) + \phi(\mathbf{e}_{k-1,k-1}) - \phi(\mathbf{e}_{k,k}) - \phi(\mathbf{e}_{k,l+1})) \\
&< 0,
\end{aligned}$$

which is a contradiction. After proceeding similarly with $\phi_{k+1,l}$, we conclude by induction that $\text{Cov}(Y_k, Y_l) - \text{Cov}(X_k, X_l) = 0$ for all $1 \leq k < l \leq d$. $\square$

Chapter 3

Supermodular ordering of binomial, Poisson and Gaussian random vectors by tree-based correlations

Summary of the Chapter

We construct a tree-based dependence structure for the representation of binomial, Poisson and Gaussian random vectors having a given covariance matrix, using sums of independent random variables. This construction allows us to characterize the supermodular ordering of such random vectors via the componentwise ordering of their covariance matrices. Our method relies on the representation of dependent components using binary trees on the discrete d -dimensional hypercube C_d , and on Möbius inversion techniques. In the case of Poisson random vectors this

approach involves Lévy measures on C_d , and it is consistent with the approximation of Poisson and multivariate Gaussian random vectors by binomial vectors.

3.1 Introduction

The supermodular stochastic ordering is used in particular to capture a preference for greater inter-dependence in economic variables.

In the case where X and Y are multivariate Gaussian vectors, supermodular ordering has been characterized by the componentwise ordering of their covariance matrices in [25] Theorem 4.2 (see Theorem 1.1.7), cf. also Theorem 3.13.5 of [26] (see Theorem 11). In [18], a dependence structure under which supermodular ordering can be characterized under the same covariance condition has been provided for Poisson random vectors, based on a decomposition of their Lévy measures on the vertices of the d -dimensional unit hypercube C_d . One can check Definition 1.1.2 for definitions of supermodular ordering and supermodular functions.

In this chapter we construct a general dependence structure for binomial and Poisson random vectors, under which the supermodular ordering can be characterized by the ordering of covariance matrices. This approach extends the results of [18] to a larger family of dependence structures, based on an arrangement of random variables according to a binary tree and on Möbius inversion. In the Gaussian case it allows us to represent any multivariate Gaussian vector using sums of independent

Gaussian random variables as in (3.2.6) below, and to recover the result of [25] as a consequence. Similarly, in the binomial, gamma and Poisson settings it yields a construction of random vectors having an arbitrarily given covariance matrix, and it provides the associated characterization of binomial and Poisson supermodular ordering, cf. Theorems 3.3.1 and 3.4.2. We refer the reader to Hu et al. [39] and references therein for the use of a different type of tree-based dependence in the setting of Bernoulli random vectors.

We proceed as follows. In Section 3.2 we construct a general dependence structure that uses binary trees on the vertices of the d -dimensional hypercube. In Section 3.3 we apply this dependence structure to the characterization of the binomial supermodular ordering via the component-wise ordering of covariances, cf. Theorem 3.3.1. In Section 3.4 we deal with the case of Poisson random vectors via the use of Lévy measures on the vertices of the d -dimensional unit hypercube C_d , cf. Theorem 3.4.2. This also includes extensions to the increasing supermodular order, cf. Proposition 3.4.3. This result naturally extends to the supermodular ordering of sums of binomial, multivariate Gaussian and Poisson random vectors. We also include a remark on the related convex ordering problem for Poisson random vectors in Proposition 3.4.4.

3.2 Tree-based correlation structures

In this section we introduce the general dependence structure used in this chapter. Let $(e_1, \dots, e_d)$ denote the canonical basis of $\mathbb{R}^d$, and let

$$C_d := \{0, 1\}^d = \{x = (x_1, \dots, x_d) : x_i \in \{0, 1\}, i = 1, \dots, d\}$$

denote the set of vertices of the d -dimensional unit hypercube.

We identify C_d to the power set $\{0, 1\}^d \simeq \{S \subset \{1, \dots, d\}\}$ of $\{1, \dots, d\}$, i.e. each $x_S = (x_1, \dots, x_d) \in C_d$ is identified to its index set $S = \{i \in \{1, \dots, d\} : x_i = 1\}$. In particular, we write $a \in x = (x_1, \dots, x_d)$ when $x_a = 1$, and $x \setminus \{a\}$ for $(x_i \mathbf{1}_{\{i \neq a\}})_{i=1, \dots, d}$.

We also endow C_d with the natural inclusion ordering of index sets i.e. we write $x \preceq y$ when $x \subseteq y$, or equivalently $0 \leq x_i \leq y_i \leq 1, i = 1, \dots, d$, and $x \prec y$ when $x \preceq y$ and $x \neq y$, i.e. $x \subsetneq y$.

Random vectors

Given $(e_{k,l})_{1 \leq k \leq l \leq d}$ a family of elements of C_d with $e_{k,k} := e_k, k = 1, \dots, d$, and a family $(X_{i,j})_{1 \leq i \leq j \leq d}$ of independent random variables, consider the random vector $X = (X_1, \dots, X_d)$ given by

$$X_i := \sum_{i \in e_{k,l}} X_{k,l}, \quad i = 1, \dots, d.$$

In other words, we have

$$X = \sum_{i=1}^d e_i \sum_{l \in e_{k,l}} X_{k,l} = \sum_{1 \leq k \leq l \leq d} X_{k,l} e_{k,l}.$$

Denoting by $(m_{k,l})_{1 \leq k \leq l \leq d}$ and $(\sigma_{k,l}^2)_{1 \leq k \leq l \leq d}$ the respective means and variances of $(X_{k,l})_{1 \leq k \leq l \leq d}$, we have

$$E[X_i] = \sum_{l \in e_{k,l}} m_{k,l}, \quad i = 1, \dots, d,$$

and

$$\text{Cov}(X_i, X_j) = \sum_{i,j \in e_{k,l}} \text{Var}[X_{k,l}] = \sum_{e_{i,j} \preceq e_{k,l}} \sigma_{k,l}^2, \quad 1 \leq i \leq j \leq d. \quad (3.2.1)$$

Binary trees

We build our tree structure in such a way that any node other than $e_{k,k} := e_k$ for $k = 1, \dots, d$ is the componentwise maximum of two sublayer nodes.

From now on we work under the following assumption on $(e_{k,l})_{1 \leq k < l \leq d}$.

- (H) The family $(e_{k,l})_{1 \leq k < l \leq d} \subset C_d$ with $e_{k,k} := e_k$, $k = 1, \dots, d$ forms a binary tree of size $d(d+1)/2$ in which every node $e_{k,l}$, $k \neq l$, has two children $e_{k,l \setminus \{k\}}$ and $e_{k,l \setminus \{l\}}$.

In particular, $(e_{k,l})_{1 \leq k < l \leq d}$ forms a binary tree with height at most d for the partial order $\preceq$.

Example. When $d = 5$, consider

CHAPTER 3. SUPERMODULAR ORDERING BY TREE-BASED CORRELATIONS

with

$$\begin{aligned}
 e_{1,2} &= \begin{vmatrix} 1 & 1 & 0 & 0 & 1 \end{vmatrix} \\
 e_{1,3} &= \begin{vmatrix} 1 & 1 & 1 & 0 & 1 \end{vmatrix} \\
 e_{1,4} &= \begin{vmatrix} 1 & 1 & 1 & 1 & 1 \end{vmatrix} \\
 e_{1,5} &= \begin{vmatrix} 1 & 0 & 0 & 0 & 1 \end{vmatrix} \\
 e_{2,3} &= \begin{vmatrix} 0 & 1 & 1 & 0 & 0 \end{vmatrix} \\
 e_{2,4} &= \begin{vmatrix} 0 & 1 & 0 & 1 & 0 \end{vmatrix} \\
 e_{2,5} &= \begin{vmatrix} 0 & 1 & 0 & 0 & 1 \end{vmatrix} \\
 e_{3,4} &= \begin{vmatrix} 0 & 1 & 1 & 1 & 0 \end{vmatrix} \\
 e_{3,5} &= \begin{vmatrix} 0 & 1 & 1 & 0 & 1 \end{vmatrix} \\
 e_{4,5} &= \begin{vmatrix} 0 & 1 & 1 & 1 & 1 \end{vmatrix}
 \end{aligned}$$

and

$$\begin{cases}
 X_1 = X_{1,1} + X_{1,2} + X_{1,3} + X_{1,4} + X_{1,5} \\
 X_2 = X_{2,2} + X_{1,2} + X_{1,3} + X_{1,4} + X_{2,3} + X_{2,4} + X_{2,5} + X_{3,4} + X_{3,5} + X_{4,5} \\
 X_3 = X_{3,3} + X_{1,3} + X_{1,4} + X_{2,3} + X_{3,4} + X_{3,5} + X_{4,5} \\
 X_4 = X_{4,4} + X_{1,4} + X_{2,4} + X_{3,4} + X_{4,5} \\
 X_5 = X_{5,5} + X_{1,2} + X_{1,3} + X_{1,4} + X_{1,5} + X_{2,5} + X_{3,5} + X_{4,5}.
 \end{cases}$$

Möbius inversion

The Möbius function is a number theoretic function which is defined by

$$\mu(n) = \begin{cases} 1 & \text{if } n \text{ is a square-free positive integer with an even number} \\ & \text{of prime factors} \\ -1 & \text{if } n \text{ is a square-free positive integer with an odd number} \\ & \text{of prime factors} \\ 0 & \text{if } n \text{ has a squared prime factor.} \end{cases}$$

The Möbius function is multiplicative function and the sum of the Möbius function over all positive divisors of n is zero except when $n = 1$. In this chapter we will use the Möbius inversion formula not the number theoretic Möbius function.

By Möbius inversion, cf. Proposition 2 of [33] (see Proposition 1.4.2) or Proposition 2.6.3 of [31], the coefficients $(\sigma_{k,l}^2)_{1 \leq k \leq l \leq d}$ in (3.2.1) can be recovered using the covariances $(\text{Cov}(X_i, X_j))_{1 \leq i \leq j \leq d}$ as

$$\sigma_{k,l}^2 = \sum_{e_{k,l} \preceq e_{i,j}} \mu(e_{i,j}, e_{k,l}) \text{Cov}(X_i, X_j), \quad 1 \leq k \leq l \leq d, \quad (3.2.2)$$

where $\mu(x, y)$ is the Möbius function defined recursively by $\mu(x, x) := 1$ and

$$\mu(x, y) = - \sum_{y \prec z \preceq x} \mu(x, z), \quad x, y \in C_d, \quad (3.2.3)$$

cf. Proposition 1 of [33] (see Proposition 1.4.1). Given $e_{k,l} \in C_d$, the two children $e_{k,l \setminus \{k\}}$, and $e_{k,l \setminus \{l\}}$ of $e_{k,l}$ have themselves a unique common

child $e_{k,l} \setminus \{k,l\}$, and (3.2.3) yields

$$\mu(e_{k,l}, e_{k,l}) = 1, \quad (3.2.4)$$

$$\mu(e_{k,l}, e_{k,l} \setminus \{k\}) = -1,$$

$$\mu(e_{k,l}, e_{k,l} \setminus \{l\}) = -1,$$

$$\mu(e_{k,l}, e_{k,l} \setminus \{k,l\}) = 1, \quad (3.2.5)$$

as shown in the next graph:

In addition, we have $\mu(e_{k,l}, e_{i,j}) = 0$ in all other cases.

In other words, (3.2.1) can be solved recursively for $(\sigma_{k,l}^2)_{1 \leq k \leq l \leq d}$ given the data of $(\text{Cov}(X_i, X_j))_{1 \leq i \leq j \leq d}$ by starting from the equality $\text{Cov}(X_k, X_l) = \sigma_{k,l}^2$ at a root $e_{k,l}$ of the tree, and then by moving down the tree step by step until each leaf e_i .

Examples. *i)* Comonotonic vectors. The binary tree contains the node $e_{k,l} = \overline{111 \cdots 11}$ as root, and all coefficients $\sigma_{i,j}^2$ vanish for $(i,j) \neq (k,l)$, which corresponds to the vector $(X_{k,l}, X_{k,l}, \dots, X_{k,l})$ with $\sigma_{k,l}^2 = \text{Var}[X_{k,l}]$.

ii) Pairwise dependence. The binary tree is reduced to the d leaves $e_1, \dots, e_d$, and to their parents

$$e_{k,l} = (0, \dots, 0, \underset{\uparrow k}{1}, 0, \dots, 0, \underset{\uparrow l}{1}, 0, \dots, 0), \quad 1 \leq k \leq l \leq d,$$

as in the following example with $d = 4$:

Here, the vector $(X_i)_{i=1,\dots,d}$ is given by

$$\begin{cases} X_1 = X_{1,1} + X_{1,2} + X_{1,3} + X_{1,4} \\ X_2 = X_{1,2} + X_{2,2} + X_{2,3} + X_{2,4} \\ X_3 = X_{1,3} + X_{2,3} + X_{3,3} + X_{3,4} \\ X_4 = X_{1,4} + X_{2,4} + X_{3,4} + X_{4,4}, \end{cases}$$

and for any $d \geq 1$, by (3.2.1) we have

$$\text{Cov}(X_i, X_j) = \sigma_{i,j}^2, \quad 1 \leq i < j \leq d,$$

and

$$\text{Var}[X_i] = \sum_{j=1}^{i-1} \sigma_{j,i}^2 + \sum_{j=i}^d \sigma_{i,j}^2, \quad i = 1, \dots, d.$$

iii) Gaussian vectors. If $(U_1, \dots, U_d)$ is a centered multivariate Gaussian random vector with covariance matrix $(\text{Cov}(U_i, U_j))_{1 \leq i \leq j \leq d}$ we can apply the Möbius inversion (3.2.2) in order to determine the variance coefficients

$(\sigma_{k,l}^2)_{1 \leq k \leq l \leq d} = (\text{Var}[Z_{k,l}])_{1 \leq k \leq l \leq d}$ of independent centered Gaussian random variables such that

$$U_i := \sum_{l \in e_{k,l}} Z_{k,l}, \quad i = 1, \dots, d. \quad (3.2.6)$$

iv) As in *(iii)* above, binomial, Poisson and gamma random vectors having a given covariance matrix can be constructed by solving (3.2.2) on a binary tree, since those distributions are characterized by their variance parameters and they are stable by summation for a given scale parameter. However in this case the construction may not be unique depending on the chosen binary tree, as their joint distribution is not characterized by their covariance matrices.

v) The dependence structure considered in [18] for Poisson random vectors corresponds to the binary tree built on the $d(d-1)/2$ nodes

$$e_{i,j} = (1, \dots, 1, \underset{\uparrow i}{1}, 0, \dots, 0, \underset{\uparrow j}{1}, 0, \dots, 0), \quad 1 \leq i < j \leq d,$$

and on the d leaves $e_1, \dots, e_d$.

3.3 Binomial random vectors

Consider $(Z_1, \dots, Z_n)$ independent Bernoulli random variables with parameter $p \in [0, 1]$ and $(A(e_{k,l}))_{1 \leq k \leq l \leq d}$ a *partition* of $\{1, \dots, n\}$. Let $(X_{A(e_{k,l})})_{1 \leq k \leq l \leq d}$ denote the family of independent binomial random vari-

ables given by

$$X_{A(e_{k,l})} := \sum_{i \in A(e_{k,l})} Z_i, \quad 1 \leq k \leq l \leq d,$$

with

$$m_{k,l} = E[X_{A(e_{k,l})}] = p|A(e_{k,l})| \quad 1 \leq k \leq l \leq d,$$

where $|A(e_{k,l})|$ denotes the cardinality of $A(e_{k,l})$, and

$$\sigma_{k,l}^2 = \text{Var}[X_{A(e_{k,l})}] = pq|A(e_{k,l})|, \quad 1 \leq k \leq l \leq d,$$

where $q := 1 - p$. Let now

$$A_i := \bigcup_{i \in e_{k,l}} A(e_{k,l}), \quad i = 1, \dots, d,$$

and consider the vector $(X_{A_1}, \dots, X_{A_d})$ of binomial random variables defined by

$$X_{A_i} := \sum_{k \in A_i} Z_k = \sum_{i \in e_{k,l}} X_{A(e_{k,l})}, \quad i = 1, \dots, d. \quad (3.3.1)$$

Example. When $d = 5$, the binary tree

CHAPTER 3. SUPERMODULAR ORDERING BY TREE-BASED CORRELATIONS

corresponds to

$$\begin{array}{rcl}
 e_{1,2} & = & \begin{vmatrix} 1 & 1 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 & 1 \end{vmatrix} \\
 e_{1,3} & = & \\
 e_{1,4} & = & \\
 e_{1,5} & = & \\
 e_{2,3} & = & \\
 e_{2,4} & = & \\
 e_{2,5} & = & \\
 e_{3,4} & = & \\
 e_{3,5} & = & \\
 e_{4,5} & = &
 \end{array}$$

and

$$\left\{ \begin{array}{l}
 X_{A_1} = X_{A(e_{1,1})} + X_{A(e_{1,2})} + X_{A(e_{1,3})} + X_{A(e_{1,4})} + X_{A(e_{1,5})} \\
 X_{A_2} = X_{A(e_{2,2})} + X_{A(e_{1,2})} + X_{A(e_{2,3})} + X_{A(e_{2,4})} + X_{A(e_{2,5})} \\
 \quad + X_{A(e_{1,3})} + X_{A(e_{3,4})} + X_{A(e_{3,5})} + X_{A(e_{4,5})} \\
 X_{A_3} = X_{A(e_{3,3})} + X_{A(e_{1,3})} + X_{A(e_{2,3})} + X_{A(e_{3,4})} + X_{A(e_{3,5})} + X_{A(e_{4,5})} \\
 X_{A_4} = X_{A(e_{4,4})} + X_{A(e_{1,4})} + X_{A(e_{2,4})} + X_{A(e_{3,4})} + X_{A(e_{4,5})} \\
 X_{A_5} = X_{A(e_{5,5})} + X_{A(e_{1,5})} + X_{A(e_{2,5})} + X_{A(e_{3,5})} + X_{A(e_{4,5})} + X_{A(e_{1,2})} + X_{A(e_{1,3})}.
 \end{array} \right.$$

In general we have

$$E[X_{A_i}] = p \sum_{i \in e_{k,l}} |A(e_{k,l})|, \quad i = 1, \dots, d,$$

and

$$\text{Cov}(X_{A_i}, X_{A_j}) = pq \sum_{e_{i,j} \preceq e_{k,l}} |A(e_{k,l})|, \quad 1 \leq i \leq j \leq d.$$

with the inversion formula

$$pq|A(e_{k,l})| = \sum_{e_{k,l} \preceq e_{i,j}} \mu(e_{i,j}, e_{k,l}) \text{Cov}(X_{A_i}, X_{A_j}), \quad 1 \leq k \leq l \leq d,$$

that follows from (3.2.2). Next is the main result of this section.

Theorem 3.3.1. *Consider $(X_{A_1}, \dots, X_{A_d})$ and $(X_{B_1}, \dots, X_{B_d})$ two binomial random vectors represented as in (3.3.1). The conditions*

$$\mathbb{E}[X_{A_i}] = \mathbb{E}[X_{B_i}], \quad 1 \leq i \leq d, \quad (3.3.2)$$

and

$$\text{Cov}(X_{A_i}, X_{A_j}) \leq \text{Cov}(X_{B_i}, X_{B_j}), \quad 1 \leq i < j \leq d, \quad (3.3.3)$$

are necessary and sufficient for the supermodular ordering

$$(X_{A_1}, \dots, X_{A_d}) \leq_{\text{sm}} (X_{B_1}, \dots, X_{B_d}).$$

Proof. It is well-known, cf. e.g. Theorem 3.9.5 of [26] (see Theorem 1.1.6), that for any couple (X, Y) of d -dimensional random vectors, the condition $X \leq_{\text{sm}} Y$ implies (3.3.2) and (3.3.3), therefore it suffices to show sufficiency. Using induction, it is also sufficient to consider the case where

$$\text{Cov}(X_{B_k}, X_{B_l}) = \text{Cov}(X_{A_k}, X_{A_l}) + pq, \quad (3.3.4)$$

for some given $1 \leq k < l \leq d$, and

$$\text{Cov}(X_{B_i}, X_{B_j}) = \text{Cov}(X_{A_i}, X_{A_j}), \quad 1 \leq i \leq j \leq d, \quad (i, j) \neq (k, l). \quad (3.3.5)$$

By the Möbius inversion formula (3.2.2) there is a unique way (up to a permutation of $\{1, \dots, n\}$) to choose $(A(e_{k,l}))_{1 \leq k \leq l \leq d}$ and $(B(e_{k,l}))_{1 \leq k \leq l \leq d}$ satisfying (3.3.4) and (3.3.5) respectively, with the relation

$$\begin{aligned} pq|B(e_{i,j})| &= \sum_{e_{i,j} \preceq e_{x,y}} \mu(e_{x,y}, e_{i,j}) \text{Cov}(X_{B_x}, X_{B_y}) \\ &= pq \mathbf{1}_{\{e_{i,j} \preceq e_{k,l}\}} \mu(e_{k,l}, e_{i,j}) + \sum_{e_{i,j} \preceq e_{x,y}} \mu(e_{x,y}, e_{i,j}) \text{Cov}(X_{A_x}, X_{A_y}) \\ &= pq \mathbf{1}_{\{e_{i,j} \preceq e_{k,l}\}} \mu(e_{k,l}, e_{i,j}) + pq|A(e_{i,j})|, \quad 1 \leq i \leq j \leq d, \end{aligned}$$

from (3.2.2), i.e.

$$|B(e_{i,j})| = \mathbf{1}_{\{e_{i,j} \preceq e_{k,l}\}} \mu(e_{k,l}, e_{i,j}) + |A(e_{i,j})|, \quad 1 \leq i \leq j \leq d. \quad (3.3.6)$$

Given the children $e_{k,l \setminus \{k\}}, e_{k,l \setminus \{l\}} \in C_d$ and grandchild $e_{k,l \setminus \{k,l\}}$ of $e_{k,l} \in C_d$, by (3.2.4)-(3.2.5) and (3.3.6) we have

$$\left\{ \begin{array}{l} |B(e_{k,l})| = |A(e_{k,l})| + 1, \\ |B(e_{k,l \setminus \{k\}})| = |A(e_{k,l \setminus \{k\}})| - 1, \\ |B(e_{k,l \setminus \{l\}})| = |A(e_{k,l \setminus \{l\}})| - 1, \\ |B(e_{k,l \setminus \{k,l\}})| = |A(e_{k,l \setminus \{k,l\}})| + 1, \end{array} \right.$$

with $|B(e_{i,j})| = |A(e_{i,j})|$ in all other cases since $\mu(e_{k,l}, e_{i,j}) = 0$. We choose to realize the above as

$$\left\{ \begin{array}{l} A(e_{k,l}) = B(e_{k,l}) \setminus \{k\}, \\ B(e_{k,l} \setminus \{k\}) = A(e_{k,l} \setminus \{k\}) \setminus \{k\}, \\ B(e_{k,l} \setminus \{l\}) = A(e_{k,l} \setminus \{l\}) \setminus \{l\}, \\ A(e_{k,l} \setminus \{k,l\}) = B(e_{k,l} \setminus \{k,l\}) \setminus \{l\}, \end{array} \right. \quad (3.3.7)$$

for some given $1 \leq k < l \leq d$, with $k, l \notin B(e_{i,j}) = A(e_{i,j})$ in all other cases. Noting that

$$l \in B(e_{k,l} \setminus \{k,l\}), \quad k \in A(e_{k,l} \setminus \{k\}), \quad l \in A(e_{k,l} \setminus \{l\}),$$

and

$$B(e_{k,l} \setminus \{k,l\}) \cap B_k = \emptyset, \quad B(e_{k,l} \setminus \{k,l\}) \cap B_l = \emptyset,$$

$$A(e_{k,l} \setminus \{k\}) \cap A_k = \emptyset, \quad A(e_{k,l} \setminus \{l\}) \cap A_l = \emptyset,$$

we find that

$$l \notin B_k, \quad l \notin B_l, \quad k \notin A_k, \quad l \notin A_l.$$

Hence, using the symmetric difference operator $A \setminus B := A \cap B^c$, for

$i = 1, \dots, d$ we have

$$A_i = \begin{cases} (B_k \setminus B(e_{k,l}) \setminus B(e_{k,l} \setminus \{k, l\})) \cup A(e_{k,l}) \cup \{l\}, & i = k, \\ (B_i \setminus B(e_{k,l}) \setminus B(e_{k,l} \setminus \{k, l\})) \cup A(e_{k,l}) \cup \{k\} \cup A(e_{k,l} \setminus \{k, l\}) \cup \{l\}, & i \notin \{k, l\}, \\ (B_l \setminus B(e_{k,l}) \setminus B(e_{k,l} \setminus \{k, l\})) \cup A(e_{k,l}) \cup \{k\}, & i = l, \end{cases} \quad (3.3.8)$$

and

$$B_i = \begin{cases} (B_k \setminus B(e_{k,l}) \setminus B(e_{k,l} \setminus \{k, l\})) \cup B(e_{k,l}), & i = k, \\ (B_i \setminus B(e_{k,l}) \setminus B(e_{k,l} \setminus \{k, l\})) \cup B(e_{k,l}) \cup B(e_{k,l} \setminus \{k, l\}), & i \notin \{k, l\}, \\ (B_l \setminus B(e_{k,l}) \setminus B(e_{k,l} \setminus \{k, l\})) \cup B(e_{k,l}), & i = l. \end{cases} \quad (3.3.9)$$

In other words, from (3.3.7) we can write

$$\begin{cases} X_{B(e_{k,l})} = X_{A(e_{k,l})} + U, \\ X_{A(e_{k,l} \setminus \{k\})} = X_{B(e_{k,l} \setminus \{k\})} + U, \\ X_{A(e_{k,l} \setminus \{l\})} = X_{B(e_{k,l} \setminus \{l\})} + V, \\ X_{B(e_{k,l} \setminus \{k, l\})} = X_{A(e_{k,l} \setminus \{k, l\})} + V, \end{cases} \quad (3.3.10)$$

where $U, V \in \{Z_1, \dots, Z_n\}$ are two independent Bernoulli random vari-

ables, while we have $X_{B(e_{i,j})} = X_{A(e_{i,j})}$ in all other cases, and from (3.3.8)-(3.3.9) we get

$$X_{A_i} = \begin{cases} X_{B_k \setminus B(e_{k,l}) \setminus B(e_{k,l} \setminus \{k,l\})} + X_{A(e_{k,l})} + V, & i = k, \\ X_{B_i \setminus B(e_{k,l}) \setminus B(e_{k,l} \setminus \{k,l\})} + X_{A(e_{k,l})} + U + X_{A(e_{k,l} \setminus \{k,l\})} + V, & i \notin \{k,l\}, \\ X_{B_l \setminus B(e_{k,l}) \setminus B(e_{k,l} \setminus \{k,l\})} + X_{A(e_{k,l})} + U, & i = l, \end{cases} \quad (3.3.11)$$

and

$$X_{B_i} = \begin{cases} X_{B_k \setminus B(e_{k,l}) \setminus B(e_{k,l} \setminus \{k,l\})} + X_{B(e_{k,l})}, & i = k, \\ X_{B_i \setminus B(e_{k,l}) \setminus B(e_{k,l} \setminus \{k,l\})} + X_{B(e_{k,l})} + X_{B(e_{k,l} \setminus \{k,l\})}, & i \notin \{k,l\}. \\ X_{B_l \setminus B(e_{k,l}) \setminus B(e_{k,l} \setminus \{k,l\})} + X_{B(e_{k,l})}, & i = l. \end{cases} \quad (3.3.12)$$

Now, for any supermodular function $\phi : \mathbb{R}^d \rightarrow \mathbb{R}$ we have, using (3.3.12) and (3.3.10),

$$\begin{aligned} & E [\phi ((X_{B_i})_{1 \leq i \leq d})] \\ &= E \left[\phi \left((X_{B_i \setminus B(e_{k,l}) \setminus B(e_{k,l} \setminus \{k,l\})} + X_{B(e_{k,l})} + X_{B(e_{k,l} \setminus \{k,l\})} \mathbf{1}_{\{i \notin \{k,l\}\}})_{1 \leq i \leq d} \right) \right] \\ &= E \left[\phi \left((X_{B_i \setminus B(e_{k,l}) \setminus B(e_{k,l} \setminus \{k,l\})} + X_{A(e_{k,l})} + U + (X_{A(e_{k,l} \setminus \{k,l\})} + V) \mathbf{1}_{\{i \notin \{k,l\}\}})_{1 \leq i \leq d} \right) \right] \\ &\geq E \left[\phi \left((X_{B_i \setminus B(e_{k,l}) \setminus B(e_{k,l} \setminus \{k,l\})} + X_{A(e_{k,l})} \right. \right. \\ &\quad \left. \left. + U \mathbf{1}_{\{i \neq k\}} + X_{A(e_{k,l} \setminus \{k,l\})} \mathbf{1}_{\{i \notin \{k,l\}\}} + V \mathbf{1}_{\{i \neq l\}})_{1 \leq i \leq d} \right) \right] \\ &= E [\phi ((X_{A_i})_{1 \leq i \leq d})], \end{aligned}$$

where we used (3.3.11) for the last equality. As for the above inequality, it follows from

$$\begin{aligned}
 & E[\phi(U, U+V, \dots, U+V, U)] \\
 &= p^2\phi(1, 2, \dots, 2, 1) + q^2\phi(0, 0, \dots, 0, 0) + pq\phi(1, 1, \dots, 1, 1) + pq\phi(0, 1, \dots, 1, 0) \\
 &\geq p^2\phi(1, 2, \dots, 2, 1) + q^2\phi(0, 0, \dots, 0, 0) + pq\phi(1, 1, \dots, 1, 0) + pq\phi(0, 1, \dots, 1, 1) \\
 &= E[\phi(U, U+V, \dots, U+V, V)],
 \end{aligned}$$

for all supermodular functions $\phi : \mathbb{R}^{|e_{k,l}|} \rightarrow \mathbb{R}$, where $|e_{k,l}|$ denotes the cardinality of $e_{k,l}$ whose indices are arranged as $\{k, \dots, l\}$ for convenience of notation, and we did not consider indices $j \notin e_{k,l}$ as U and V do not belong to X_j in this case. $\square$

Multivariate Gaussian vectors

From the central limit theorem, Theorem 3.3.1 can be used to deal with centered multivariate Gaussian random vectors $(U_1, \dots, U_d)$ and $(V_1, \dots, V_d)$ with covariance matrices

$$(\text{Cov}(U_i, U_j))_{1 \leq i \leq j \leq d} \quad \text{and} \quad (\text{Cov}(V_i, V_j))_{1 \leq i \leq j \leq d}.$$

In this case we can apply the Möbius inversion (3.2.2) in order to determine the variance coefficients $(\sigma_{k,l}^2)_{1 \leq k \leq l \leq d}$ and $(\eta_{k,l}^2)_{1 \leq k \leq l \leq d}$ in the decomposition (3.2.6). Those coefficients can then be obtained as the respective limits of variances $(\text{Var}[X_{k,l}^n/\sqrt{n}])_{1 \leq k \leq l \leq d}$ and $(\text{Var}[Y_{k,l}^n/\sqrt{n}])_{1 \leq k \leq l \leq d}$ of in-

dependent binomial random variables such that

$$U_i^n := \frac{1}{\sqrt{n}} \sum_{i \in e_{k,l}} (X_{k,l}^n - E[X_{k,l}^n]) \quad \text{and} \quad V_i^n := \frac{1}{\sqrt{n}} \sum_{i \in e_{k,l}} (Y_{k,l}^n - E[Y_{k,l}^n]), \quad i = 1, \dots, d,$$

converge in distribution to $(U_1, \dots, U_d)$ and $(V_1, \dots, V_d)$ respectively. Letting n tend to infinity in $\text{Cov}(U_i^n, U_j^n) \leq \text{Cov}(V_i^n, V_j^n)$, the condition $\text{Cov}(U_i, U_j) \leq \text{Cov}(V_i, V_j)$ of Theorem 3.3.1, $1 \leq i < j \leq d$, becomes necessary and sufficient for $(U_1, \dots, U_d) \leq_{\text{sm}} (V_1, \dots, V_d)$ to hold. In this way we recover the result of [25], Theorem 4.2 (see Theorem 1.1.7), for Gaussian random vectors, cf. also Theorem 3.13.5 of [26] (see Theorem 1.1.7).

A similar argument results from Theorem 3.4.2 below in the Poisson case, using the convergence in distribution from renormalized binomial random variables to Poisson random variables. In the next section we provide a proof of such a result using Lévy measures for infinitely divisible Poisson random vectors.

3.4 Poisson random vectors

Recall that any d -dimensional infinitely divisible Poisson random vector $X = (X_1, \dots, X_d)$ is defined by its characteristic function

$$E[e^{i\langle \bar{t}, X \rangle}] = \exp \left(\int_{\mathbb{R}^d} (e^{i\langle \bar{t}, x \rangle} - 1) \mu(dx) \right),$$

where $\bar{t} = (t_1, \dots, t_d) \in \mathbb{R}^d$, $\langle \cdot, \cdot \rangle$ denotes the scalar product in $\mathbb{R}^d$, and the Lévy measure

$$\mu(dx) := \sum_{\emptyset \neq S \subset \{1, 2, \dots, d\}} a_S \delta_{e_S}(dx),$$

is supported on C_d , where $\delta_{e_S}(dx)$ denotes the Dirac measure at the point $e_S \in C_d$, and $(a_S)_{\emptyset \neq S \subset \{1, 2, \dots, d\}}$ is a family of nonnegative coefficients.

Equivalently, $X = (X_1, \dots, X_d)$ can be represented as

$$X_i = \sum_{\substack{S \in \{0, 1\}^d \\ S \neq \emptyset}} \mathbf{1}_{\{i \in S\}} X_S = \sum_{\substack{S \subset \{1, 2, \dots, d\} \\ S \ni i}} X_S, \quad i = 1, \dots, d,$$

where $(X_S)_{\emptyset \neq S \subset \{1, 2, \dots, d\}}$ is a family of $2^d - 1$ independent Poisson random variables with respective intensities $(a_S)_{\emptyset \neq S \subset \{1, 2, \dots, d\}}$, cf. also Theorem 3 of [16] (see Theorem 1.2.1).

In order to characterize the ordering of Poisson random vectors based on the data of their covariance matrices which contain only $d(d+1)/2$ components, we restrict ourselves to Lévy measures of the form

$$\mu(dx) = \sum_{1 \leq k \leq l \leq d} a_{k,l} \delta_{e_{k,l}}(dx), \quad (3.4.1)$$

on C_d , where $a_{k,l} \in \mathbb{R}_+$, $1 \leq k \leq l \leq d$. In other words, we have

$$X = \sum_{i=1}^d e_i \sum_{i \in e_{k,l}} X_{k,l} = \sum_{1 \leq k \leq l \leq d} X_{k,l} e_{k,l}, \quad (3.4.2)$$

where $(X_{k,l})_{1 \leq k \leq l \leq d}$ is a family of independent Poisson random variables whose respective intensity parameters $(a_{i,j})_{1 \leq i \leq j \leq d}$ satisfy $\text{Var}[X_{k,l}] = E[X_{k,l}] = a_{k,l}$, $1 \leq k \leq l \leq d$, with the inversion formula

$$a_{k,l} = \sum_{e_{k,l} \preceq e_{i,j}} \mu(e_{i,j}, e_{k,l}) \text{Cov}(X_i, X_j), \quad 1 \leq k \leq l \leq d, \quad (3.4.3)$$

that follows from (3.2.2).

Supermodular ordering of Poisson random vectors

Theorem 3.4.2 below is a direct consequence of the following Lemma 3.4.1 which provides the decomposition

$$\mu(dx) = \sum_{i=1}^d \text{Var}[X_i] \delta_{e_i} + \sum_{1 \leq i < j \leq d} \text{Cov}(X_i, X_j) (\delta_{e_{i,j}} + \delta_{e_{i,j} \setminus \{i,j\}} - \delta_{e_{i,j} \setminus \{i\}} - \delta_{e_{i,j} \setminus \{j\}})$$

of the Lévy measure $\mu(dx)$ on $C_d \setminus \{0\}$, using the covariance matrix of $(X_i)_{i=1,\dots,d}$.

Lemma 3.4.1. *For any function $\phi : C_d \rightarrow \mathbb{R}$ such that $\phi(0) = 0$ we have*

$$\begin{aligned} \int_{\mathbb{R}^d} \phi(x) \mu(dx) &= \sum_{i=1}^d \mathbb{E}[X_i] \phi(e_i) \\ &+ \sum_{1 \leq i < j \leq d} \text{Cov}(X_i, X_j) (\phi(e_{i,j}) + \phi(e_{i,j} \setminus \{i,j\}) - \phi(e_{i,j} \setminus \{i\}) - \phi(e_{i,j} \setminus \{j\})). \end{aligned}$$

Proof. By the Möbius inversion formula (3.2.2) we have

$$\int_{\mathbb{R}^d} \phi(x) \mu(dx) = \sum_{1 \leq k \leq l \leq d} a_{k,l} \phi(e_{k,l})$$

$$\begin{aligned}
&= \sum_{1 \leq k \leq l \leq d} \phi(e_{k,l}) \sum_{e_{k,l} \preceq e_{i,j}} \mu(e_{i,j}, e_{k,l}) \text{Cov}(X_i, X_j) \\
&= \sum_{i=1}^d \text{Cov}(X_i, X_i) \sum_{e_k \preceq e_i} \mu(e_i, e_k) \phi(e_k) + \sum_{1 \leq i < j \leq d} \text{Cov}(X_i, X_j) \sum_{\substack{e_{k,l} \preceq e_{i,j} \\ 1 \leq k < l \leq d}} \mu(e_{i,j}, e_{k,l}) \phi(e_{k,l}) \\
&= \sum_{i=1}^d E[X_i] \phi(e_i) + \sum_{1 \leq i < j \leq d} \text{Cov}(X_i, X_j) (\phi(e_{i,j}) + \phi(e_{i,j} \setminus \{i, j\}) - \phi(e_{i,j} \setminus \{i\}) - \phi(e_{i,j} \setminus \{j\})),
\end{aligned}$$

where we used (3.2.4)-(3.2.5) and the fact that $e_k \preceq e_i$ if and only if $k = i$. $\square$

Consider now two Poisson random vectors X and Y whose respective Lévy measures μ and ν are represented as

$$\mu(dx) = \sum_{1 \leq i \leq j \leq d} a_{i,j} \delta_{e_{i,j}}(dx) \quad \text{and} \quad \nu(dx) = \sum_{1 \leq i \leq j \leq d} b_{i,j} \delta_{e_{i,j}}(dx),$$

as in (3.4.1). If X_i has the same distribution as Y_i for all $i = 1, \dots, d$ then $E[X_i] = E[Y_i]$, $i = 1, \dots, d$, and Lemma 3.4.1 shows that

$$\begin{aligned}
&\int_{\mathbb{R}^d} \phi(y) \nu(dy) - \int_{\mathbb{R}^d} \phi(x) \mu(dx) \\
&= \sum_{1 \leq i < j \leq d} (\text{Cov}(Y_i, Y_j) - \text{Cov}(X_i, X_j)) (\phi(e_{i,j}) + \phi(e_{i,j} \setminus \{i, j\}) - \phi(e_{i,j} \setminus \{i\}) - \phi(e_{i,j} \setminus \{j\})).
\end{aligned} \tag{3.4.4}$$

Relation (3.4.4) shows in particular that the nonnegativity of the coefficients

$$\text{Cov}(Y_i, Y_j) - \text{Cov}(X_i, X_j) \geq 0, \quad 1 \leq i < j \leq d, \tag{3.4.5}$$

becomes a necessary and sufficient condition for the supermodular ordering of the (finite support) Lévy measures μ and ν .

The next Theorem 3.4.2 reformulates (3.4.5) as a necessary and sufficient condition for supermodular ordering of infinitely divisible Poisson random vector, based on Theorem 4.5 of [3] (see Theorem 1.3.3) which allows us to carry over the notion of supermodularity from the finite support setting of Lévy measures μ, ν on the cube C_d , to the infinite support setting of Poisson random variables.

Theorem 3.4.2. *Consider two Poisson random vectors X and Y both represented as in (3.4.2). Then the conditions*

$$\mathbb{E}[X_i] = \mathbb{E}[Y_i], \quad 1 \leq i \leq d,$$

and

$$\text{Cov}(X_i, X_j) \leq \text{Cov}(Y_i, Y_j), \quad 1 \leq i < j \leq d, \quad (3.4.6)$$

are necessary and sufficient for the supermodular ordering $X \leq_{\text{sm}} Y$.

Proof. By Theorem 4.5 in [3] (see Theorem 1.3.3) it suffices to show that

$$\int_{\mathbb{R}^d} \phi(x) \mu(dx) \leq \int_{\mathbb{R}^d} \phi(y) \nu(dy) \quad (3.4.7)$$

for all supermodular functions $\phi : \mathbb{R}^d \rightarrow \mathbb{R}$, where $\mu(dx)$ and $\nu(dy)$ respectively denote the Lévy measures of X and Y . By Lemma 3.4.1 we have the identity

$$\begin{aligned} & \int_{\mathbb{R}^d} \phi(y) \nu(dy) - \int_{\mathbb{R}^d} \phi(x) \mu(dx) \\ &= \sum_{1 \leq i < j \leq d} (\text{Cov}(Y_i, Y_j) - \text{Cov}(X_i, X_j)) (\phi(e_{i,j}) + \phi(e_{i,j \setminus \{i,j\}}) - \phi(e_{i,j \setminus \{i\}}) - \phi(e_{i,j \setminus \{j\}})) \end{aligned}$$

under condition (3.4.6), which allows us to conclude to (3.4.7) for all supermodular functions ϕ . $\square$

The next proposition is obtained as in Proposition 4.3 of [18] (see Proposition 2.4.3) by extending Theorem 4.5 of [3] (see Theorem 1.3.3) to nondecreasing supermodular functions ϕ on $\mathbb{R}^d$ satisfying $\phi(0) = 0$, using the same approximation as in Lemma 4.4 (see Lemma 1.3.2) therein.

Proposition 3.4.3. *Consider two Poisson random vectors X and Y both represented as in (3.4.2), and assume that*

$$\mathbb{E}[X_i] \leq \mathbb{E}[Y_i], \quad 1 \leq i \leq d,$$

and

$$\text{Cov}(X_i, X_j) \leq \text{Cov}(Y_i, Y_j), \quad 1 \leq i < j \leq d.$$

Then we have

$$\mathbb{E}[\Phi(X)] \leq \mathbb{E}[\Phi(Y)]$$

for all nondecreasing supermodular functions $\Phi : \mathbb{R}^d \longrightarrow \mathbb{R}$.

Sums of binomial, Gaussian and Poisson vectors

By Theorem 4.2 (see Theorem 1.1.7) of [25] on Gaussian random vectors, Theorems 3.3.1 and 3.4.2 above, and the fact that the supermodular ordering is closed under convolution, cf. Theorem 3.9.14-(C) of [26] (see Theorem 1.1.4), we deduce that the supermodular ordering of a sum of independent binomial, Gaussian and Poisson vectors, is implied by the componentwise ordering of their respective covariances. Proposition 3.4.3

admits an analog extension to sums of binomial, Gaussian and Poisson random vectors.

Convex ordering

Proposition 3.4.4. *Consider two Poisson random vectors X and Y both represented as in (3.4.2). Then we have $X \leq_{\text{cx}} Y$ if and only if X and Y have same distributions.*

Proof. Assume $X \leq_{\text{cx}} Y$, i.e. we have

$$\mathbb{E}[\Phi(X)] \leq \mathbb{E}[\Phi(Y)]$$

for all *convex* functions $\Phi : \mathbb{R}^d \rightarrow \mathbb{R}$. Clearly, this implies $E[X_k] = E[Y_k]$, $k = 1, \dots, d$. Next, choosing any $1 \leq k < l \leq d$ we check that the function

$$(x_1, \dots, x_d) \mapsto \phi_{k,l}(x_1, \dots, x_d) := \max \left(0, x_l - x_k - \sum_{a \notin e_{k,l}} x_a \right)$$

is convex on $\mathbb{R}^d$, with $\phi_{k,l}(e_{i,j}) = 1$ when $e_{i,j}$ is a (non-strict) descendant of $e_{k,l} \setminus \{k\}$ that contains l , and $\phi_{k,l}(e_{i,j}) = 0$ in all other cases. This yields

$$\phi_{k,l}(e_{k,l}) + \phi_{k,l}(e_{k,l} \setminus \{k, l\}) - \phi_{k,l}(e_{k,l} \setminus \{k\}) - \phi_{k,l}(e_{k,l} \setminus \{l\}) = -1,$$

and

$$\phi_{k,l}(e_{i,j}) + \phi_{k,l}(e_{i,j} \setminus \{i, j\}) - \phi_{k,l}(e_{i,j} \setminus \{i\}) - \phi_{k,l}(e_{i,j} \setminus \{j\}) = 0$$

when $(i, j) \neq (k, l)$. Hence by Lemma 3.4.1, the condition $\text{Cov}(Y_k, Y_l) <$

$\text{Cov}(X_k, X_l)$ would imply

$$\begin{aligned}
& \int_{\mathbb{R}^d} \phi(y) \nu(dy) - \int_{\mathbb{R}^d} \phi(x) \mu(dx) \\
&= \sum_{1 \leq i < j \leq d} (\text{Cov}(Y_i, Y_j) - \text{Cov}(X_i, X_j)) (\phi(e_{i,j}) + \phi(e_{i,j} \setminus \{i,j\}) - \phi(e_{i,j} \setminus \{i\}) - \phi(e_{i,j} \setminus \{j\})) \\
&= (\text{Cov}(Y_k, Y_l) - \text{Cov}(X_k, X_l)) (\phi(e_{k,l}) + \phi(e_{k,l} \setminus \{k,l\}) - \phi(e_{k,l} \setminus \{k\}) - \phi(e_{k,l} \setminus \{l\})) \\
&> 0,
\end{aligned}$$

which is equivalent to $X \leq_{\text{cx}} Y$ by the same argument as in part (b) of the proof of Theorem 4.5 in [3] (see Theorem 1.3.3).

Proceeding similarly by taking the convex function

$$(x_1, \dots, x_d) \mapsto \phi_{k,l}(x_1, \dots, x_d) := \max \left(1, x_l + x_k - \sum_{a \notin e_{k,l}} x_a \right)$$

we will have

$$\phi_{k,l}(e_{k,l}) + \phi_{k,l}(e_{k,l} \setminus \{k,l\}) - \phi_{k,l}(e_{k,l} \setminus \{k\}) - \phi_{k,l}(e_{k,l} \setminus \{l\}) = 1,$$

and

$$\phi_{k,l}(e_{i,j}) + \phi_{k,l}(e_{i,j} \setminus \{i,j\}) - \phi_{k,l}(e_{i,j} \setminus \{i\}) - \phi_{k,l}(e_{i,j} \setminus \{j\}) = 0$$

when $(i,j) \neq (k,l)$. Hence by Lemma 3.4.1, the condition $\text{Cov}(Y_k, Y_l) \geq \text{Cov}(X_k, X_l)$ would imply

$$\begin{aligned}
& \int_{\mathbb{R}^d} \phi(y) \nu(dy) - \int_{\mathbb{R}^d} \phi(x) \mu(dx) \\
&= \sum_{1 \leq i < j \leq d} (\text{Cov}(Y_i, Y_j) - \text{Cov}(X_i, X_j)) (\phi(e_{i,j}) + \phi(e_{i,j} \setminus \{i,j\}) - \phi(e_{i,j} \setminus \{i\}) - \phi(e_{i,j} \setminus \{j\})) \\
&= (\text{Cov}(Y_k, Y_l) - \text{Cov}(X_k, X_l)) (\phi(e_{k,l}) + \phi(e_{k,l} \setminus \{k,l\}) - \phi(e_{k,l} \setminus \{k\}) - \phi(e_{k,l} \setminus \{l\}))
\end{aligned}$$

$$\geq 0,$$

which is equivalent to $X \leq_{\text{cx}} Y$. Then we deduce that $\text{Cov}(Y_k, Y_l) = \text{Cov}(X_k, X_l)$ for all $1 \leq k < l \leq d$, hence X and Y have same distribution from (3.4.3). $\square$

3.5 Some counter examples

In this section we state two counter examples of 4-dimension which do not satisfy the conditions of Theorem 3.4.2. Under these considered structures of Poisson random vectors; we will see in the following examples necessary and sufficient conditions of Theorem 3.4.2 will not imply supermodular ordering of Poisson random vectors.

Example 3.5.1. Let X_{ij} be independent Poisson random variables with intensities $a_{ij} \in \mathbb{R}_+$ for $1 \leq i \leq j \leq 4$. Assume that we have the following dependence structure

$$\begin{cases} X_1 = X_{1,1} + X_{1,2} + X_{1,3} + X_{1,4} + X_{2,3} \\ X_2 = X_{1,2} + X_{2,2} + X_{2,3} + X_{2,4} + X_{3,4} \\ X_3 = X_{1,3} + X_{2,3} + X_{3,3} + X_{3,4} \\ X_4 = X_{1,4} + X_{2,4} + X_{3,4} + X_{4,4} \end{cases} . \quad (3.5.1)$$

Also let Y_{ij} be independent Poisson random variables with intensities $b_{ij} \in \mathbb{R}_+$ for $1 \leq i \leq j \leq 4$ and assume that we have the following

dependence structure

$$\begin{cases} Y_1 = Y_{1,1} + Y_{1,2} + Y_{1,3} + Y_{1,4} + Y_{2,3} \\ Y_2 = Y_{1,2} + Y_{2,2} + Y_{2,3} + Y_{2,4} + Y_{3,4} \\ Y_3 = Y_{1,3} + Y_{2,3} + Y_{3,3} + Y_{3,4} \\ Y_4 = Y_{1,4} + Y_{2,4} + Y_{3,4} + Y_{4,4} \end{cases} . \quad (3.5.2)$$

For the structure (3.5.1) we will have

$$\begin{aligned} & \int_{\mathbb{R}^4} \phi(x) \mu(dx) \\ &= \sum_{i=1}^4 \mathbb{E}[X_i] \phi(e_i) \\ &+ \text{Cov}(X_1, X_2) (\phi(1, 1, 0, 0) + \phi(0, 0, 0, 0) - \phi(1, 0, 0, 0) - \phi(0, 1, 0, 0)) \\ &+ \text{Cov}(X_1, X_3) (\phi(1, 0, 1, 0) + \phi(0, 0, 0, 0) - \phi(1, 0, 0, 0) - \phi(0, 0, 1, 0)) \\ &+ \text{Cov}(X_1, X_4) (\phi(1, 0, 0, 1) + \phi(0, 0, 0, 0) - \phi(1, 0, 0, 0) - \phi(0, 0, 0, 1)) \\ &+ \text{Cov}(X_2, X_3) (\phi(1, 1, 1, 0) + \phi(1, 0, 0, 0) - \phi(1, 1, 0, 0) - \phi(1, 0, 1, 0)) \\ &+ \text{Cov}(X_2, X_4) (\phi(0, 1, 0, 1) + \phi(0, 0, 0, 0) - \phi(0, 1, 0, 0) - \phi(0, 0, 0, 1)) \\ &+ \text{Cov}(X_3, X_4) (\phi(0, 1, 1, 1) + \phi(0, 0, 0, 0) + \phi(1, 0, 1, 0) + \phi(1, 1, 0, 0) \\ &\quad - \phi(0, 1, 0, 1) - \phi(1, 1, 1, 0) - \phi(0, 0, 1, 0) - \phi(1, 0, 0, 0)) \end{aligned}$$

and for the structure (3.5.2) we will have

$$\begin{aligned} & \int_{\mathbb{R}^4} \phi(x) \nu(dx) \\ &= \sum_{i=1}^4 \mathbb{E}[Y_i] \phi(e_i) \\ &+ \text{Cov}(Y_1, Y_2) (\phi(1, 1, 0, 0) + \phi(0, 0, 0, 0) - \phi(1, 0, 0, 0) - \phi(0, 1, 0, 0)) \end{aligned}$$

$$\begin{aligned}
& + \text{Cov}(Y_1, Y_3)(\phi(1, 0, 1, 0) + \phi(0, 0, 0, 0) - \phi(1, 0, 0, 0) - \phi(0, 0, 1, 0)) \\
& + \text{Cov}(Y_1, Y_4)(\phi(1, 0, 0, 1) + \phi(0, 0, 0, 0) - \phi(1, 0, 0, 0) - \phi(0, 0, 0, 1)) \\
& + \text{Cov}(Y_2, Y_3)(\phi(1, 1, 1, 0) + \phi(1, 0, 0, 0) - \phi(1, 1, 0, 0) - \phi(1, 0, 1, 0)) \\
& + \text{Cov}(Y_2, Y_4)(\phi(0, 1, 0, 1) + \phi(0, 0, 0, 0) - \phi(0, 1, 0, 0) - \phi(0, 0, 0, 1)) \\
& + \text{Cov}(Y_3, Y_4)(\phi(0, 1, 1, 1) + \phi(0, 0, 0, 0) + \phi(1, 0, 1, 0) + \phi(1, 1, 0, 0) \\
& - \phi(0, 1, 0, 1) - \phi(1, 1, 1, 0) - \phi(0, 0, 1, 0) - \phi(1, 0, 0, 0)).
\end{aligned}$$

Under the following conditions:

i) $\mathbb{E}[X_i] = \mathbb{E}[Y_i], \quad 1 \leq i \leq 4$

ii) $\text{Cov}(X_i, X_j) \leq \text{Cov}(Y_i, Y_j), \quad 1 \leq i < j \leq 4,$

we may not have $\int_{\mathbb{R}^4} \phi(x) \mu(dx) \leq \int_{\mathbb{R}^4} \phi(x) \nu(dx)$, i.e. $X \leq_{\text{sm}} Y$.

The reason is that for all ϕ supermodular functions we have

$$\begin{aligned}
& \int_{\mathbb{R}^4} \phi(x) \nu(dx) - \int_{\mathbb{R}^4} \phi(x) \mu(dx) \\
& = \sum_{\substack{1 \leq i < j \leq 4 \\ (i,j) \neq (3,4)}} (\text{Cov}(Y_i, Y_j) - \text{Cov}(X_i, X_j)) (\phi(e_{i,j}) + \phi(e_{i,j} \setminus \{i, j\}) - \phi(e_{i,j} \setminus \{i\}) - \phi(e_{i,j} \setminus \{j\})) \\
& + (\text{Cov}(Y_3, Y_4) - \text{Cov}(X_3, X_4)) (\phi(0, 1, 1, 1) + \phi(0, 0, 0, 0) + \phi(1, 0, 1, 0) + \phi(1, 1, 0, 0) \\
& - \phi(0, 1, 0, 1) - \phi(1, 1, 1, 0) - \phi(0, 0, 1, 0) - \phi(1, 0, 0, 0))
\end{aligned}$$

hence the last term may not be positive when ϕ is supermodular, for example, by taking

$$\phi(x_1, x_2, x_3, x_4) = x_1 x_2 x_3, \quad x_i \in \{0, 1\}, \quad 1 \leq i \leq 4$$

we will have

$$\begin{aligned}
 & \phi(0, 1, 1, 1) + \phi(0, 0, 0, 0) + \phi(1, 0, 1, 0) + \phi(1, 1, 0, 0) \\
 & - \phi(0, 1, 0, 1) - \phi(1, 1, 1, 0) - \phi(0, 0, 1, 0) - \phi(1, 0, 0, 0) \\
 & = -1 < 0.
 \end{aligned}$$

Then we cannot compare $\int_{\mathbb{R}^4} \phi(x) \mu(dx)$ and $\int_{\mathbb{R}^4} \phi(x) \nu(dx)$ under conditions (i) and (ii) above.

In other way, for the following dependence structure

$$\begin{cases}
 X_1 = X_{1,1} + X_{1,2} + X_{1,3} + X_{1,4} + X_{2,3} \\
 X_2 = X_{1,2} + X_{2,2} + X_{2,3} + X_{2,4} + X_{3,4} \\
 X_3 = X_{1,3} + X_{2,3} + X_{3,3} + X_{3,4} \\
 X_4 = X_{1,4} + X_{2,4} + X_{3,4} + X_{4,4}
 \end{cases}$$

we will have the following table

$$\begin{array}{rcl}
 e_{1,2} & = & \begin{vmatrix} 1 & 1 & 0 & 0 \end{vmatrix} \\
 e_{1,3} & = & \begin{vmatrix} 1 & 0 & 1 & 0 \end{vmatrix} \\
 e_{1,4} & = & \begin{vmatrix} 1 & 0 & 0 & 1 \end{vmatrix} \\
 e_{2,3} & = & \begin{vmatrix} 1 & 1 & 1 & 0 \end{vmatrix} \\
 e_{2,4} & = & \begin{vmatrix} 0 & 1 & 0 & 1 \end{vmatrix} \\
 e_{3,4} & = & \begin{vmatrix} 0 & 1 & 1 & 1 \end{vmatrix}
 \end{array}$$

and for this table we will have the following tree

In above tree, the node '0111' stands alone however according to our tree structure rule (H) this node should be the componentwise maximum of two sub layer nodes. Since it does not then we cannot compare the above dependence structure with respect to supermodular order.

Example 3.5.2. Let X_{ij} be independent Poisson random variables with intensities $a_{ij} \in \mathbb{R}_+$ for $1 \leq i \leq j \leq 4$. Assume that we have the following dependence structure

$$\begin{cases} X_1 = X_{1,1} + X_{1,2} + X_{1,3} + X_{1,4} + X_{2,4} + X_{3,4} \\ X_2 = X_{1,2} + X_{2,2} + X_{2,3} + X_{2,4} + X_{3,4} \\ X_3 = X_{1,3} + X_{2,3} + X_{3,3} + X_{3,4} \\ X_4 = X_{1,4} + X_{2,4} + X_{3,4} + X_{4,4} \end{cases} . \quad (3.5.3)$$

Also let Y_{ij} be independent Poisson random variables with intensities $b_{ij} \in \mathbb{R}_+$ for $1 \leq i \leq j \leq 4$ and assume that we have the following

dependence structure

$$\begin{cases} Y_1 = Y_{1,1} + Y_{1,2} + Y_{1,3} + Y_{1,4} + Y_{2,4} + Y_{3,4} \\ Y_2 = Y_{1,2} + Y_{2,2} + Y_{2,3} + Y_{2,4} + Y_{3,4} \\ Y_3 = Y_{1,3} + Y_{2,3} + Y_{3,3} + Y_{3,4} \\ Y_4 = Y_{1,4} + Y_{2,4} + Y_{3,4} + Y_{4,4} \end{cases} . \quad (3.5.4)$$

For the structure (3.5.3) we will have

$$\begin{aligned} & \int_{\mathbb{R}^4} \phi(x) \mu(dx) \\ &= \sum_{i=1}^4 \mathbb{E}[X_i] \phi(e_i) \\ &+ \text{Cov}(X_1, X_2) (\phi(1, 1, 0, 0) + \phi(0, 0, 0, 0) - \phi(1, 0, 0, 0) - \phi(0, 1, 0, 0)) \\ &+ \text{Cov}(X_1, X_3) (\phi(1, 0, 1, 0) + \phi(0, 0, 0, 0) - \phi(1, 0, 0, 0) - \phi(0, 0, 1, 0)) \\ &+ \text{Cov}(X_1, X_4) (\phi(1, 0, 0, 1) + \phi(0, 0, 0, 0) - \phi(1, 0, 0, 0) - \phi(0, 0, 0, 1)) \\ &+ \text{Cov}(X_2, X_3) (\phi(0, 1, 1, 0) + \phi(0, 0, 0, 0) - \phi(0, 1, 0, 0) - \phi(0, 0, 1, 0)) \\ &+ \text{Cov}(X_2, X_4) (\phi(1, 1, 0, 1) + \phi(1, 0, 0, 0) - \phi(1, 1, 0, 0) - \phi(1, 0, 0, 1)) \\ &+ \text{Cov}(X_3, X_4) (\phi(1, 1, 1, 1) + \phi(1, 0, 0, 0) + \phi(0, 1, 0, 0) + \phi(0, 0, 1, 0) \\ &\quad - \phi(1, 0, 1, 0) - \phi(0, 0, 0, 0) - \phi(0, 1, 1, 0) - \phi(1, 1, 0, 1)) \end{aligned}$$

and for the structure (3.5.4) we will have

$$\begin{aligned} & \int_{\mathbb{R}^4} \phi(x) \nu(dx) \\ &= \sum_{i=1}^4 \mathbb{E}[Y_i] \phi(e_i) \\ &+ \text{Cov}(Y_1, Y_2) (\phi(1, 1, 0, 0) + \phi(0, 0, 0, 0) - \phi(1, 0, 0, 0) - \phi(0, 1, 0, 0)) \end{aligned}$$

$$\begin{aligned}
& + \text{Cov}(Y_1, Y_3)(\phi(1, 0, 1, 0) + \phi(0, 0, 0, 0) - \phi(1, 0, 0, 0) - \phi(0, 0, 1, 0)) \\
& + \text{Cov}(Y_1, Y_4)(\phi(1, 0, 0, 1) + \phi(0, 0, 0, 0) - \phi(1, 0, 0, 0) - \phi(0, 0, 0, 1)) \\
& + \text{Cov}(Y_2, Y_3)(\phi(0, 1, 1, 0) + \phi(0, 0, 0, 0) - \phi(0, 1, 0, 0) - \phi(0, 0, 1, 0)) \\
& + \text{Cov}(Y_2, Y_4)(\phi(1, 1, 0, 1) + \phi(1, 0, 0, 0) - \phi(1, 1, 0, 0) - \phi(1, 0, 0, 1)) \\
& + \text{Cov}(Y_3, Y_4)(\phi(1, 1, 1, 1) + \phi(1, 0, 0, 0) + \phi(0, 1, 0, 0) + \phi(0, 0, 1, 0) \\
& - \phi(1, 0, 1, 0) - \phi(0, 0, 0, 0) - \phi(0, 1, 1, 0) - \phi(1, 1, 0, 1)).
\end{aligned}$$

Under the following conditions:

i) $\mathbb{E}[X_i] = \mathbb{E}[Y_i], \quad 1 \leq i \leq 4$

ii) $\text{Cov}(X_i, X_j) \leq \text{Cov}(Y_i, Y_j), \quad 1 \leq i < j \leq 4$

we may not have $\int_{\mathbb{R}^4} \phi(x) \mu(dx) \leq \int_{\mathbb{R}^4} \phi(x) \nu(dx)$, i.e. $X \leq_{\text{sm}} Y$.

The reason is that for all ϕ supermodular functions we have

$$\begin{aligned}
& \int_{\mathbb{R}^4} \phi(x) \nu(dx) - \int_{\mathbb{R}^4} \phi(x) \mu(dx) \\
& = \sum_{\substack{1 \leq i < j \leq 4 \\ (i,j) \neq (3,4)}} (\text{Cov}(Y_i, Y_j) - \text{Cov}(X_i, X_j)) (\phi(e_{i,j}) + \phi(e_{i,j} \setminus \{i, j\}) - \phi(e_{i,j} \setminus \{i\}) - \phi(e_{i,j} \setminus \{j\})) \\
& + (\text{Cov}(Y_3, Y_4) - \text{Cov}(X_3, X_4)) (\phi(1, 1, 1, 1) + \phi(1, 0, 0, 0) + \phi(0, 1, 0, 0) + \phi(0, 0, 1, 0) \\
& - \phi(1, 0, 1, 0) - \phi(0, 0, 0, 0) - \phi(0, 1, 1, 0) - \phi(1, 1, 0, 1)).
\end{aligned}$$

hence the last term may not be positive when ϕ is supermodular, for example, by taking

$$\phi(x_1, x_2, x_3, x_4) = \exp(-x_2 x_4 - x_3 x_4) - 1, \quad x_i \in \{0, 1\}, \quad 1 \leq i \leq 4$$

we will have

$$\begin{aligned}
 & \phi(1, 1, 1, 1) + \phi(1, 0, 0, 0) + \phi(0, 1, 0, 0) + \phi(0, 0, 1, 0) \\
 & - \phi(1, 0, 1, 0) - \phi(0, 0, 0, 0) - \phi(0, 1, 1, 0) - \phi(1, 1, 0, 1) \\
 & = e^{-2} - e^{-1} < 0.
 \end{aligned}$$

Then we cannot compare $\int_{\mathbb{R}^4} \phi(x) \mu(dx)$ and $\int_{\mathbb{R}^4} \phi(x) \nu(dx)$ under conditions (i) and (ii) above.

In other way, for the following dependence structure

$$\begin{cases} X_1 = X_{1,1} + X_{1,2} + X_{1,3} + X_{1,4} + X_{2,4} + X_{3,4} \\ X_2 = X_{1,2} + X_{2,2} + X_{2,3} + X_{2,4} + X_{3,4} \\ X_3 = X_{1,3} + X_{2,3} + X_{3,3} + X_{3,4} \\ X_4 = X_{1,4} + X_{2,4} + X_{3,4} + X_{4,4} \end{cases}$$

we will have the following array table

$$\begin{array}{lcl} e_{1,2} & = & \begin{vmatrix} 1 & 1 & 0 & 0 \end{vmatrix} \\ e_{1,3} & = & \begin{vmatrix} 1 & 0 & 1 & 0 \end{vmatrix} \\ e_{1,4} & = & \begin{vmatrix} 1 & 0 & 0 & 1 \end{vmatrix} \\ e_{2,3} & = & \begin{vmatrix} 0 & 1 & 1 & 0 \end{vmatrix} \\ e_{2,4} & = & \begin{vmatrix} 1 & 1 & 0 & 1 \end{vmatrix} \\ e_{3,4} & = & \begin{vmatrix} 1 & 1 & 1 & 1 \end{vmatrix} \end{array} .$$

and for this array table we will have the following tree

CHAPTER 3. SUPERMODULAR ORDERING BY TREE-BASED CORRELATIONS

In above tree, the node '1111' stands alone however according to our tree structure rule (H) this node should be the componentwise maximum of two sub layer nodes. Since it does not then we cannot compare the above dependence structure with respect to supermodular order.

Chapter 4

Bounds for option prices by supermodular ordering

4.1 Introduction

Ordering and comparison of stochastic models are topics of particular interest which have undergone an intensive development. The main aim is to derive approximation results and bounds for complex stochastic models where fundamental model assumptions are only partially known.

Consider the following multidimensional drift-diffusion-jump model

$$dX(t) = \mu(X(t), t)dt + \beta(X(t), t)dW(t) + \int_0^1 \phi(X(t), t, z)\tilde{\nu}(dt, dz) \quad (4.1.1)$$

where $X(t)$ is a d -dimensional stochastic process, $\mu = (\mu_1, \dots, \mu_d)$ is a d -dimensional vector, $\beta = (\beta_{ij})_{i,j=1}^d$ is a $d \times d$ -matrix, $W(t)$ is a d -dimensional Brownian motion, $\phi = (\phi_1, \dots, \phi_d)$ is a d -dimensional vec-

tor, and $\tilde{\nu}$ is the compensated jump measure defined by $\tilde{\nu}(dt, dz) = (\nu - q)(dt, dz)$ such that ν is a Poisson random measure on $[0, T] \times [0, 1]$ with intensity measure $q(dt, dz) = \lambda(t)dt dz$ where λ is a deterministic function.

In following we state the definition of supermodularity preserving model. One can find definitions of supermodular ordering and supermodular functions in Definition 1.1.2.

Definition 4.1.1. *Let $f : \mathbb{R}^d \rightarrow \mathbb{R}$ be a supermodular function and define the option price at time $t \in [0, T]$ by*

$$F(t, x) = \mathbb{E}[f(X(T)) | X(t) = x] \quad (4.1.2)$$

then we call the stochastic model (4.1.1) a 'supermodularity preserving model' if F is a supermodular function.

Our main aim is to find conditions on μ , β , and ϕ in (4.1.1) so that our price function $F : [0, T] \times \mathbb{R}_+^d \rightarrow \mathbb{R}$ will be a supermodular function.

In [5] and [4] sufficient conditions for the comparison of European options with respect to martingale pricing measures are derived. Their comparison is with respect to convex and directionally convex functions. They formulate sufficient conditions in terms of predictable characteristics of the stochastic logarithm of the stock price processes. They show that ordering of the differential characteristics of the stochastic logarithms X, X^* of two exponential semimartingale models S and S^* should imply ordering of the prices p, p^* of European options with convex payoff function. In order to establish this conclusion for general models, a technical

condition needs to be satisfied, the "propagation of convexity property". They develop a new approach to establish this property for several examples of univariate and multivariate processes. For example, they provide comparison results for exponential semimartingales, multivariate diffusion processes, stochastic volatility models, Lévy processes, and diffusion with jumps.

It has been shown in [3] that comparing two Lévy processes with respect to supermodular ordering is equivalent to comparing their Lévy measures with respect to supermodular ordering. By using this amazing result, in [17] we construct a general dependence structure for binomial and Poisson random vectors, by using binary trees and Möbius inversion techniques, to show that supermodular ordering of binomial and Poisson random vectors can be characterized by ordering of their covariance matrices, and this is an extension of results in [18] to a much larger family.

The case of convex ordering has been considered by Ekström and Tysk [10] in jump-diffusion models. They showed that the only higher-dimensional convexity preserving models are the ones with linear coefficients. They obtained necessary and sufficient conditions on the coefficients of a parabolic equation for preservation of convexity. In multidimensional case, under appropriate growth conditions, they showed that the solution of a parabolic equation remains convex if the initial condition is a convex function.

In this chapter, we prove that multidimensional Ornstein-Uhlenbeck process and multidimensional jump-diffusion model are supermodularity preserving models and then by using results in [17] we obtain sufficient conditions for nontrivial bounds of option prices for multidimensional

Ornstein-Uhlenbeck process, and in a similar way one can derive sufficient conditions for bounds of option prices for multidimensional jump-diffusion model.

We proceed in the following way: in Section 4.2 we prove the propagation of supermodularity for multidimensional Ornstein-Uhlenbeck process and then we derive sufficient conditions to obtain nontrivial bounds for option prices. In Section 4.4 we prove propagation of supermodularity property for multidimensional jump-diffusion model. In Section 4.5 we define locally supermodularity preserving (LSP) operators and then we show that if our model is supermodularity preserving then it is LSP.

4.2 Preservation of supermodularity for multidimensional Ornstein-Uhlenbeck process

A multidimensional Ornstein-Uhlenbeck process X_t^* defined on a probability space $(\Omega, \mathcal{F}^*, (\mathcal{F}_t^*)_{t \in \mathbb{R}_+}, \mathbb{P}^*)$ satisfies the following linear stochastic differential equation (SDE) on $\mathbb{R}^d$:

$$dX_t^* = (A_t^* X_t^* + \alpha_t^*)dt + B_t^* dW_t^* + C_t^* dL_t^*, \quad X_0^* = x_0^* \quad (4.2.1)$$

where $t \rightarrow A_t^*$, $t \rightarrow B_t^*$, and $t \rightarrow C_t^*$ are bounded deterministic functions in the space of $d \times d$ matrices and $t \rightarrow \alpha_t^*$ is a bounded deterministic function in the space of d -dimensional vectors, W_t^* is a d -dimensional Wiener process, and L_t^* is a Poisson process with measure ν^* on $\mathbb{R}^d$.

Explicit solution of (4.2.1) is given in the following way:

$$X_t^* = \Psi_t \left(x_0^* + \int_0^t \Psi_s^{-1} \alpha_s^* ds + \int_0^t \Psi_s^{-1} B_s^* dW_s^* + \int_0^t \Psi_s^{-1} C_s^* dL_s^* \right) \quad (4.2.2)$$

for $t \geq 0$. Here Ψ_t is a $d \times d$ matrix at time t with $\Psi_0 = I$, where I denotes the identity matrix and it is explicitly given in the form

$$\Psi_t = \exp \left(\int_0^t A_s^* ds \right), \quad (4.2.3)$$

where the exponential is a matrix exponential.

In the following lemma, we prove ‘the propagation of supermodularity’ property for multidimensional Ornstein-Uhlenbeck process.

Lemma 4.2.1. *Let $f : \mathbb{R}^d \rightarrow \mathbb{R}$ be a twice differentiable supermodular function and X_t^* be a multidimensional Ornstein-Uhlenbeck process defined as in (4.2.1). If A_t^* is a $d \times d$ matrix with nonnegative entries then*

$$F(t, x) = \mathbb{E}^*[f(X_T^*) | X_t^* = x] \quad (\in C^{1,2}([0, T] \times \mathbb{R}^d)) \quad (4.2.4)$$

is a supermodular function. Moreover, $F(t, X_t^)$ is an $(\mathcal{F}_t^*)_{t \in \mathbb{R}_+}$ -martingale.*

Proof. We have

$$\begin{aligned} F(t, x) &= \mathbb{E}^*[f(X_T^*) | X_t^* = x] \\ &= \mathbb{E}^* \left[f \left(\Psi_T \left(X_t^* + \int_t^T \Psi_s^{-1} \alpha_s^* ds + \int_t^T \Psi_s^{-1} B_s^* dW_s^* + \int_t^T \Psi_s^{-1} C_s^* dL_s^* \right) \right) \middle| X_t^* = x \right] \\ &= \mathbb{E}^* \left[f \left(\Psi_T \left(x + \int_t^T \Psi_s^{-1} \alpha_s^* ds + \int_t^T \Psi_s^{-1} B_s^* dW_s^* + \int_t^T \Psi_s^{-1} C_s^* dL_s^* \right) \right) \right]. \end{aligned}$$

By taking partial derivative of $F(t, x)$ with respect to x_i we will have

$$\begin{aligned} & \frac{\partial F}{\partial x_i}(t, x) \\ &= \mathbb{E}^* \left[\sum_{k=1}^d (\Psi_T)_{ik} \frac{\partial f}{\partial x_i} \left(\Psi_T \left(x + \int_t^T \Psi_s^{-1} \alpha_s^* ds + \int_t^T \Psi_s^{-1} B_s^* dW_s^* + \int_t^T \Psi_s^{-1} C_s^* dL_s^* \right) \right) \right]. \end{aligned} \quad (4.2.5)$$

Then by taking partial derivative of (4.2.5) with respect to x_j ($i \neq j$) we will have

$$\begin{aligned} \frac{\partial^2 F}{\partial x_i \partial x_j}(t, x) &= \mathbb{E}^* \left[\sum_{k=1}^d (\Psi_T)_{ik} \sum_{k=1}^d (\Psi_T)_{jk} \right. \\ &\quad \times \frac{\partial^2 f}{\partial x_i \partial x_j} \left(\Psi_T \left(x + \int_t^T \Psi_s^{-1} \alpha_s^* ds + \int_t^T \Psi_s^{-1} B_s^* dW_s^* + \int_t^T \Psi_s^{-1} C_s^* dL_s^* \right) \right) \Big]. \end{aligned}$$

By using supermodularity of f and the assumption that A_t^* has nonnegative entries; then we will have

$$\frac{\partial^2 F}{\partial x_i \partial x_j}(t, x) \geq 0.$$

□

4.3 Bounds of option prices for multidimensional Ornstein-Uhlenbeck process

In this section, our aim is to find sufficient conditions for comparing two multidimensional processes with respect to increasing supermodular ordering. We bound option prices by using multidimensional Ornstein-

Uhlenbeck process.

Assume that X_t is a multidimensional stochastic process defined on $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in \mathbb{R}_+}, \mathbb{P})$ and given by:

$$dX_t = a(t, X_t)dt + B_t dW_t + C_t dL_t \quad (4.3.1)$$

where $t \rightarrow B_t$ and $t \rightarrow C_t$ are bounded deterministic functions in the space of $d \times d$ matrices, $x \rightarrow a(t, x)$ ($[0, \infty) \times \mathbb{R}^d \rightarrow \mathbb{R}^d$) is Lipschitz uniform; W_t is a d -dimensional Brownian motion; and L_t is a Poisson process with measure ν on $\mathbb{R}^d$.

Also assume that the Poisson measures of (4.2.1) and (4.3.1) are defined on the unit hypercube of $\mathbb{R}^d$:

$$C_d = \{0, 1\}^d = \left\{ e_S := \sum_{i \in S} e_i : S \subset \{1, \dots, d\} \right\}$$

which take forms

$$\nu^*(dx) = \sum_{\emptyset \neq S \subset \{1, 2, \dots, d\}} a_S \delta_{e_S}(dx),$$

and

$$\nu(dx) = \sum_{\emptyset \neq S \subset \{1, 2, \dots, d\}} b_S \delta_{e_S}(dx),$$

where $(a_S)_{\emptyset \neq S \subset \{1, 2, \dots, d\}}, (b_S)_{\emptyset \neq S \subset \{1, 2, \dots, d\}}$ are families of nonnegative numbers. Moreover we specialize our $\nu^*(dx)$ and $\nu(dx)$ as in (4.1) of [17] (see 3.4.1).

In order to prove our main result we use the Kolmogorov backward equation and it is given in the following lemma:

Lemma 4.3.1. *Let X_t^* be a multidimensional Ornstein-Uhlenbeck process as in (4.2.1). Then $F(t, x) = \mathbb{E}^*[f(X_T^*)|X_t^* = x]$ satisfies the Kolmogorov backward equation:*

$$\begin{aligned} \frac{\partial F}{\partial t}(t, X_t^*) = & - \sum_{i \leq d} \frac{\partial F}{\partial x_i}(t, X_t^*)(A_t^* X_t^* + \alpha_t^*)^i - \frac{1}{2} \sum_{i, j \leq d} \frac{\partial^2 F}{\partial x_i \partial x_j}(t, X_t^*)(B_t^*)^{ij} \\ & - \int_{C_d} (F(t, X_t^* + C_t^* x) - F(t, X_t^*)) \nu^*(dx). \end{aligned} \quad (4.3.2)$$

Proof. We follow a similar strategy as in the proof of Lemma 2.1 [4].

From Itô's formula we have

$$\begin{aligned} F(t, X_t^*) = & F(0, X_0^*) + \sum_{i \leq d} \int_{[0, t]} \frac{\partial F}{\partial x_i}(s, X_s^*)(B_s^*)^i dW_s^{*i} \\ & + \int_{[0, t]} \int_{C_d} (F(s, X_s^* + C_s^* x) - F(s, X_s^*)) (dL_s^* - \nu^*(dx) ds) \\ & + \int_{[0, t]} \left(\frac{\partial F}{\partial s}(s, X_s^*) + \sum_{i \leq d} \frac{\partial F}{\partial x_i}(s, X_s^*)(A_s^* X_s^* + \alpha_s^*)^i \right. \\ & \left. + \frac{1}{2} \sum_{i, j \leq d} \frac{\partial^2 F}{\partial x_i \partial x_j}(s, X_s^*)(B_s^*)^{ij} + \int_{C_d} (F(s, X_s^* + C_s^* x) - F(s, X_s^*)) \nu^*(dx) \right) ds. \end{aligned}$$

Since $F(t, X_t^*)$ is an $\mathcal{F}_t^*$ -martingale then we have

$$\begin{aligned} \frac{\partial F}{\partial t}(t, X_t^*) = & - \sum_{i \leq d} \frac{\partial F}{\partial x_i}(t, X_t^*)(A_t^* X_t^* + \alpha_t^*)^i - \frac{1}{2} \sum_{i, j \leq d} \frac{\partial^2 F}{\partial x_i \partial x_j}(t, X_t^*)(B_t^*)^{ij} \\ & - \int_{C_d} (F(t, X_t^* + C_t^* x) - F(t, X_t^*)) \nu^*(dx). \end{aligned}$$

□

The following theorem is the main result of this chapter; we state sufficient conditions for comparison of two multidimensional processes

with respect to increasing supermodular order. We use multidimensional Ornstein-Uhlenbeck process which is supermodularity preserving to bound a multidimensional process whose analytic solution is not known. Indeed, we are trying to find nontrivial bounds for option prices.

Theorem 4.3.2. *Let X_t^* (Ornstein-Uhlenbeck) and X_t be two multidimensional processes whose stochastic differential equations are given as in (4.2.1) and (4.3.1), respectively. Also assume that A_t^*, C_t^* , and C_t are $d \times d$ matrices with nonnegative entries with $C_t^* = C_t$. Under following conditions*

$$(a(t, x))^i \leq (A_t^* x + \alpha_t^*)^i, \quad x \in \mathbb{R}^d, \quad i \in \{1, 2, \dots, d\}, \quad (4.3.3)$$

$$(B_t)^{ij} \leq (B_t^*)^{ij}, \quad i, j \in \{1, 2, \dots, d\}, \quad (4.3.4)$$

$$\text{Cov}(L_t^i, L_t^j) \leq \text{Cov}(L_t^{*i}, L_t^{*j}), \quad i, j \in \{1, 2, \dots, d\} \quad (4.3.5)$$

we will have

$$\mathbb{E}[f(X_T)] \leq \mathbb{E}^*[f(X_T^*)]$$

for all f increasing supermodular functions.

Proof. From Itô's formula we have

$$\begin{aligned} F(t, X_t) = & F(0, X_0) + \sum_{i \leq d} \int_{[0, t]} \frac{\partial F}{\partial x_i}(s, X_s) (B_s)^i dW_s^i \\ & + \int_{[0, t]} \int_{C_d} (F(s, X_s + C_s x) - F(s, X_s)) (dL_s - \nu(dx) ds) \\ & + \int_{[0, t]} \left(\frac{\partial F}{\partial s}(s, X_s) + \sum_{i \leq d} \frac{\partial F}{\partial x_i}(s, X_s) a(s, X_s)^i \right. \end{aligned}$$

$$+\frac{1}{2} \sum_{i,j \leq d} \frac{\partial^2 F}{\partial x_i \partial x_j}(s, X_s)(B_s)^{ij} + \int_{C_d} (F(s, X_s + C_s x) - F(s, X_s)) \nu(dx) \Big) ds. \quad (4.3.6)$$

If we insert the Kolmogorov backward equation:

$$\begin{aligned} \frac{\partial F}{\partial s}(s, X_s) = & - \sum_{i \leq d} \frac{\partial F}{\partial x_i}(s, X_s)(A_s^* X_s + \alpha_s^*)^i - \frac{1}{2} \sum_{i,j \leq d} \frac{\partial^2 F}{\partial x_i \partial x_j}(s, X_s)(B_s^*)^{ij} \\ & - \int_{C_d} (F(s, X_s + C_s^* x) - F(s, X_s)) \nu^*(dx) \end{aligned} \quad (4.3.7)$$

into (4.3.6) we will have

$$\begin{aligned} F(t, X_t) = & F(0, X_0) + \sum_{i \leq d} \int_{[0,t]} \frac{\partial F}{\partial x_i}(s, X_s)(B_s)^i dW_s^i \\ & + \int_{[0,t]} \int_{C_d} (F(s, X_s + C_s x) - F(s, X_s)) (dL_s - \nu(dx) ds) \\ & + \int_{[0,t]} \sum_{i \leq d} \frac{\partial F}{\partial x_i}(s, X_s) (a(s, X_s)^i - (A_s^* X_s + \alpha_s^*)^i) ds \\ & + \int_{[0,t]} \frac{1}{2} \sum_{i,j \leq d} \frac{\partial^2 F}{\partial x_i \partial x_j}(s, X_s) ((B_s)^{ij} - (B_s^*)^{ij}) ds \\ & + \int_{[0,t]} \int_{C_d} (F(s, X_s + C_s x) - F(s, X_s)) (\nu(dx) - \nu^*(dx)) ds. \end{aligned} \quad (4.3.8)$$

Since $F(t, \cdot)$ is an increasing supermodular function and since C_t is a $d \times d$ matrix with nonnegative entries then

$$x \longmapsto F(s, X_s + C_s x) - F(s, X_s)$$

is also a supermodular function. Moreover

$$\frac{\partial F}{\partial x_i}, \frac{\partial^2 F}{\partial x_i \partial x_j}, \text{ and } F(s, X_s + C_s x) - F(s, X_s)$$

are nonnegative quantities.

It has been shown in Theorem 4.5 of [3] (see Theorem 1.3.3) that comparing two Lévy processes with respect to supermodular ordering is equivalent to comparing their Lévy measures with respect to supermodular ordering. Moreover in [18] (see Theorem 2.4.2) and [17] (see Theorem 3.4.2) we showed that comparing two Poisson random vectors with respect to supermodular ordering is equivalent to comparing their covariance matrices componentwise. Therefore, instead of comparing

$$\nu \leq_{\text{sm}} \nu^*$$

i.e.

$$\begin{aligned} & \int_{[0,t]} \int_{C_d} (F(s, X_s + C_s x) - F(s, X_s)) \nu(dx) ds \\ & \leq \int_{[0,t]} \int_{C_d} (F(s, X_s + C_s x) - F(s, X_s)) \nu^*(dx) ds \end{aligned}$$

for $F(s, X_s + C_s x) - F(s, X_s)$ supermodular function, we will compare

$$\text{Cov}(L_t^i, L_t^j) \leq \text{Cov}(L_t^{*i}, L_t^{*j}), \text{ for all } i, j \in \{1, 2, \dots, d\}.$$

Finally, if we take expectations on both sides of (4.3.8) and use given conditions then we will have

$$\mathbb{E}[F(t, X_t)] \leq F(0, X_0)$$

i.e.

$$\mathbb{E}[F(t, X_t)] = \mathbb{E}[\mathbb{E}^*[f(X_T^*)|X_t^* = X_t]] \leq F(0, X_0) = \mathbb{E}^*[f(X_T^*)|X_0^* = X_0]$$

and finally we will have

$$\mathbb{E}[f(X_T)] \leq \mathbb{E}^*[f(X_T^*)].$$

□

4.4 Preservation of supermodularity for multidimensional jump-diffusion model

A multidimensional jump-diffusion model, $X_t = (X_1(t), \dots, X_d(t))$, satisfies the following linear stochastic differential equations:

$$dX_i(t) = r(t)X_i(t)dt + X_i(t) \sum_{j=1}^d \sigma_{ij}(t)dW_j(t) + X_i(t) \int_0^1 \gamma_i(t, z)\tilde{\nu}(dt, dz), \quad i = 1, 2, \dots, d, \quad (4.4.1)$$

where $t \rightarrow r(t)$, $t \rightarrow \sigma_{ij}(t)$ are deterministic functions, $W_j(t)$ is a Brownian motion, $t \rightarrow \gamma_i(t, z)$ is a deterministic function for a fixed $z \in [0, 1]$ and $\tilde{\nu}$ is the compensated jump measure defined by $\tilde{\nu}(dt, dz) = (\nu - q)(dt, dz)$ such that ν is a Poisson random measure on $[0, T] \times [0, 1]$ with intensity measure $q(dt, dz) = \lambda(t)dt dz$ where λ is a deterministic function.

The solution of (4.4.1) is given as follows

$$X_i(t) = X_i(0) \exp \left(\int_0^t \left(r(s) - \frac{1}{2} \sum_{j=1}^d \sigma_{ij}^2(s) \right) ds + \sum_{j=1}^d \int_0^t \sigma_{ij}(s) dW_j(s) \right) J_i(t) \quad (4.4.2)$$

where

$$J_i(t) = \exp \left(- \int_0^t \int_0^1 \lambda(s) \gamma_i(s, z) dz ds \right) \prod_{0 \leq z \leq 1} \prod_{0 \leq \tau \leq t} (1 + \gamma_i(\tau))^{\nu(z, \tau)}. \quad (4.4.3)$$

The preservation of supermodularity property for multidimensional jump-diffusion model is given in the following lemma:

Lemma 4.4.1. *Let $f : \mathbb{R}^d \rightarrow \mathbb{R}$ be a twice differentiable supermodular function and let $X_t = (X_1(t), \dots, X_d(t))$ be a multidimensional jump-diffusion model as in (4.4.1) with $\gamma_i > -1$ for $1 \leq i \leq d$. Then*

$$F(t, x) = \mathbb{E}[f(X_T) | X_t = x] \quad (F : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}) \quad (4.4.4)$$

is a supermodular function.

Proof. By taking partial derivatives of $F(t, x)$ with respect to x_i and x_j for $i \neq j$ we will have

$$\begin{aligned} & \frac{\partial^2 F}{\partial x_i \partial x_j}(t, x) \\ &= J_i(T) J_j(T) \exp \left(\int_t^T \left(2r(s) - \frac{1}{2} \sum_{k=1}^d (\sigma_{ik}^2(s) + \sigma_{jk}^2(s)) \right) ds \right) \\ & \times \mathbb{E} \left[\exp \left(\sum_{k=1}^d \int_t^T (\sigma_{ik}(s) + \sigma_{jk}(s)) dW_k(s) \right) \frac{\partial^2 f}{\partial x_i \partial x_j}(\dots) \right] \geq 0 \end{aligned}$$

for all f supermodular functions and $\gamma_i > -1$. □

4.5 Locally supermodularity preserving operators

In this section, first, we give conditions for the existence and uniqueness of the solution of (4.1.1). Second, we define locally supermodularity preserving operators. Third, we show that if a stochastic model is supermodularity preserving then the sum of operators which consist of drift, diffusion, and jump parts is locally supermodularity preserving. Lastly, we give a necessary and sufficient condition for the differential operator which represent the jump part to be supermodularity preserving.

Under the following regularity and growth assumptions (see for example Theorem 3.1 of Mao [19] and Pages 717-718 of Ekström and Tysk [10]) there is a unique solution of (4.1.1): for a positive constant D and Hölder exponent $\alpha \in (0, 1)$ if we have

A1) $|\mu(x, t) - \mu(y, t)|^2 \bigvee |\beta(x, t) - \beta(y, t)|^2 \leq D|x - y|^2$ for all $x, y \in \mathbb{R}^d$ and $t \in [0, T]$;

A2) $|\mu(x, t)|^2 \bigvee |\beta(x, t)|^2 \leq D(1 + |x|^2)$ for all $(x, t) \in \mathbb{R}^d \times [0, T]$;

A3) The matrix $\beta(x, t)$ is non-singular for all $(x, t) \in \mathbb{R}_+^d \times [0, T]$;

A4) For all $i, j = 1, \dots, d$, $\beta_{ij} : \mathbb{R}_+ \times [0, T] \rightarrow \mathbb{R}$ is in $C_\alpha^{2,0}(\mathbb{R}_+^d \times [0, T])$;

A5) $|\beta(x, t) - \beta(y, t)| + |\phi(x, t, z) - \phi(y, t, z)| \leq D|x - y|$;

A6) $\sum_{j=1}^d |\beta_{ij}(x, t)|^2 + |\phi_i(x, t, z)|^2 \leq Dx_i^2$ for all $i = 1, \dots, d$;

A7) $\phi : \mathbb{R}_+^d \times [0, T] \times [0, 1] \rightarrow \mathbb{R}^d$ is measurable, and $\phi_i(\cdot, \cdot, z) \in C_\alpha^{2,0}(\mathbb{R}_+^d \times [0, T])$ with the Hölder norms being uniformly bounded in the

z -variable, also for all x and t we have $\phi(x, t, z) \neq 0$ for almost all z ;

A8) There exists $\gamma > -1$ such that $\phi_i(x, t, z) > \gamma x_i$ for $i = 1, \dots, d$;

A9) $\lambda \in C_\alpha([0, T])$.

By $C_{pol}(\mathbb{R}_+ \times [0, T])$ we denote the set of functions of at most polynomial growth in x . More explicitly,

$$C_{pol}(\mathbb{R} \times [0, T]) := \bigcup_{C>0, m>0} \{f : \mathbb{R} \times [0, T] \rightarrow \mathbb{R} : |f(x, t)| \leq C(1 + |x|^m) \text{ for } (x, t) \in \mathbb{R} \times [0, T]\}.$$

By $C_\alpha(\mathbb{R}_+ \times [0, T])$ we denote the set of locally Hölder functions, i.e.

$$C_\alpha(\mathbb{R}_+ \times [0, T]) := \left\{ f : \mathbb{R}_+ \times [0, T] \rightarrow \mathbb{R} : \sup_{p, q \in K} \frac{|f(p) - f(q)|}{d(p, q)^\alpha} < \infty \text{ for each compact } K \subset \mathbb{R}_+ \times [0, T] \right\}$$

where $d((x, t), (y, s)) = (|x - y|^2 + |t - s|)^{1/2}$.

We define locally supermodularity preserving operators in the following way:

Definition 4.5.1. *An operator $\mathcal{M}$ is called locally supermodularity preserving (LSP) at a point $(x, t) \in \mathbb{R}_+^d \times [0, T]$ if for any two different $i, j \in \{1, 2, \dots, d\}$ we have*

$$\frac{\partial^2(\mathcal{M}f)}{\partial x_i \partial x_j}(x, t) \geq 0 \tag{4.5.1}$$

for all supermodular functions $f \in C_\alpha^4(\mathbb{R}_+^d) \cap C_{pol}^2(\mathbb{R}_+^d)$ with $\frac{\partial^2 f}{\partial x_i \partial x_j}(x) = 0$.

Consider the following parabolic integro-differential equation

$$\frac{\partial \mathcal{F}}{\partial t}(x, t) + (\mathcal{A} + \mathcal{B} + \mathcal{C})\mathcal{F}(x, t) = 0, \quad \mathcal{F}(x, T) = f(x). \quad (4.5.2)$$

where the first-order differential operator $\mathcal{A}$, the second-order differential operator $\mathcal{B}$ and the integro-differential operator $\mathcal{C}$ are defined by

$$\begin{aligned} \mathcal{A}\mathcal{F}(x, t) &:= \sum_{i=1}^d \mu_i(x, t) \frac{\partial \mathcal{F}}{\partial x_i}(x, t); \\ \mathcal{B}\mathcal{F}(x, t) &:= \frac{1}{2} \sum_{i,j=1}^d a_{ij}(x, t) \frac{\partial^2 \mathcal{F}}{\partial x_i \partial x_j}(x, t); \end{aligned}$$

and

$$\mathcal{C}\mathcal{F}(x, t) := \lambda(t) \int_0^1 (\mathcal{F}(x + \phi(x, t, z), t) - \mathcal{F}(x, t) - \phi(x, t, z) \cdot \nabla \mathcal{F}(x, t)) dz$$

respectively, where a_{ij} for $1 \leq i, j \leq d$ are coefficients of the matrix $\beta\beta^T/2$.

Then we have the following main result of this section:

Theorem 4.5.2. *If the multidimensional drift-diffusion-jump model (4.1.1) is supermodularity preserving then $\mathcal{A} + \mathcal{B} + \mathcal{C}$ is locally supermodularity preserving (LSP) at all points $(x, t) \in \mathbb{R}_+^d \times [0, T]$.*

Proof. We will follow the same strategy as in Theorem 4.2 in [10]. Choose $(x_0, T_0) \in \mathbb{R}_+^d \times [0, T]$ and $f \in C_\alpha^4(\mathbb{R}_+^d) \cap C_{pol}^2(\mathbb{R}_+^d)$ be a supermodular function such that $\frac{\partial^2 f}{\partial x_i \partial x_j}(x_0) = 0$.

Consider the solution $\mathcal{F}(x, t)$ of

$$\frac{\partial \mathcal{F}}{\partial t}(x, t) + (\mathcal{A} + \mathcal{B} + \mathcal{C})\mathcal{F}(x, t) = 0 \quad \text{with } \mathcal{F}(x, T) = f(x).$$

Also let $\tilde{f} \in C_\alpha^4(\mathbb{R}_+^d)$ be a supermodular, Lipschitz continuous function which satisfies $\tilde{f} = f$ in a domain which contains x_0 and all possible values of $x_0 + \phi$. Then we have $\mathcal{A}\tilde{f} = \mathcal{A}f$, $\mathcal{B}\tilde{f} = \mathcal{B}f$, and $\mathcal{C}\tilde{f} = \mathcal{C}f$ in a (spatial) neighborhood of x_0 . So

$$\frac{\partial^2(\mathcal{A} + \mathcal{B} + \mathcal{C})\tilde{f}}{\partial x_i \partial x_j}(x_0, T_0) = \frac{\partial^2(\mathcal{A} + \mathcal{B} + \mathcal{C})f}{\partial x_i \partial x_j}(x_0, T_0). \quad (4.5.3)$$

Consider the solution $\tilde{\mathcal{F}}(x, t)$ of

$$\frac{\partial \tilde{\mathcal{F}}}{\partial t}(x, t) + (\mathcal{A} + \mathcal{B} + \mathcal{C})\tilde{\mathcal{F}}(x, t) = 0 \quad \text{with } \tilde{\mathcal{F}}(x, T) = \tilde{f}(x). \quad (4.5.4)$$

Since the model is supermodularity preserving, $\tilde{\mathcal{F}}$ is supermodular in x at all times prior to T_0 , and in particular $\frac{\partial^2 \tilde{\mathcal{F}}}{\partial x_i \partial x_j}(x_0, t) \geq 0$.

Since

$$\frac{\partial^2 \tilde{\mathcal{F}}}{\partial x_i \partial x_j}(x_0, T_0) = \frac{\partial^2 \tilde{f}}{\partial x_i \partial x_j}(x_0) = \frac{\partial^2 f}{\partial x_i \partial x_j}(x_0) = 0$$

thus $\frac{\partial^3 \tilde{\mathcal{F}}}{\partial t \partial x_i \partial x_j}(x_0, t)$ exist and by above argument we have $\frac{\partial^3 \tilde{\mathcal{F}}}{\partial t \partial x_i \partial x_j}(x_0, T_0) \leq 0$.

Then by using (4.5.3) and (4.5.4) we will have

$$\begin{aligned} 0 &\leq -\frac{\partial^3 \tilde{\mathcal{F}}}{\partial t \partial x_i \partial x_j}(x_0, T_0) \\ &= -\frac{\partial^2}{\partial x_i \partial x_j} \left(\frac{\partial \tilde{\mathcal{F}}}{\partial t}(x_0, T_0) \right) \\ &= \frac{\partial^2}{\partial x_i \partial x_j} \left((\mathcal{A} + \mathcal{B} + \mathcal{C})\tilde{\mathcal{F}}(x_0, T_0) \right) \\ &= \frac{\partial^2}{\partial x_i \partial x_j} \left((\mathcal{A} + \mathcal{B} + \mathcal{C})\tilde{f}(x_0) \right) \end{aligned}$$

$$= \frac{\partial^2}{\partial x_i \partial x_j} ((\mathcal{A} + \mathcal{B} + \mathcal{C})f(x_0)).$$

In conclusion, $\mathcal{A} + \mathcal{B} + \mathcal{C}$ is locally supermodularity preserving (LSP) at $(x_0, T_0) \in \mathbb{R}_+^d \times [0, T]$. Since $(x_0, T_0) \in \mathbb{R}_+^d \times [0, T]$ is arbitrary then $\mathcal{A} + \mathcal{B} + \mathcal{C}$ is LSP at all points $(x, t) \in \mathbb{R}_+^d \times [0, T]$. $\square$

In the following proposition we give a necessary and sufficient condition for the operator $\mathcal{C}$ to be locally supermodularity preserving.

Proposition 4.5.3. *The operator $\mathcal{C}$ is locally supermodularity preserving (LSP) at a point $(x, t) \in \mathbb{R}_+^d \times [0, T]$ if and only if*

$$\begin{aligned} & \int_0^1 \left(\left(1 + \frac{\partial \phi_i}{\partial x_i}(x, t, z)\right) \left(1 + \frac{\partial \phi_j}{\partial x_j}(x, t, z)\right) \frac{\partial^2 f}{\partial x_i \partial x_j}(x + \phi(x, t, z)) \right. \\ & \quad + \frac{\partial^2 \phi_i}{\partial x_i \partial x_j}(x, t, z) \frac{\partial f}{\partial x_i}(x + \phi(x, t, z)) - \sum_{k=1}^d \frac{\partial^2 \phi_k}{\partial x_i \partial x_j}(x, t, z) \frac{\partial f}{\partial x_k}(x) \\ & \quad \left. - \frac{\partial \phi_i}{\partial x_j}(x, t, z) \frac{\partial^2 f}{\partial x_i^2}(x) - \frac{\partial \phi_j}{\partial x_i}(x, t, z) \frac{\partial^2 f}{\partial x_j^2}(x) \right) dz \geq 0, \end{aligned} \quad (4.5.5)$$

for all supermodular functions $f \in C_\alpha^4(\mathbb{R}_+^d) \cap C_{pol}^2(\mathbb{R}_+^d)$ with $\frac{\partial^2 f}{\partial x_i \partial x_j}(x) = 0$, for $1 \leq i < j \leq d$.

Proof. For $f : \mathbb{R}_+^d \rightarrow \mathbb{R}$ and $\phi : \mathbb{R}_+^d \times [0, T] \times [0, 1] \rightarrow \mathbb{R}^d$ we have

$$\begin{aligned} & \phi(x, t, z) \cdot \nabla f(x) \\ &= (\phi_1(x, t, z), \phi_2(x, t, z), \dots, \phi_d(x, t, z)) \cdot \left(\frac{\partial f}{\partial x_1}(x), \frac{\partial f}{\partial x_2}(x), \dots, \frac{\partial f}{\partial x_d}(x) \right) \\ &= \phi_1(x, t, z) \frac{\partial f}{\partial x_1}(x) + \phi_2(x, t, z) \frac{\partial f}{\partial x_2}(x) + \dots + \phi_d(x, t, z) \frac{\partial f}{\partial x_d}(x). \end{aligned}$$

After taking partial derivatives of

$$\mathcal{C}f(x, t) = \lambda(t) \int_0^1 (f(x + \phi(x, t, z)) - f(x) - \phi(x, t, z) \cdot \nabla f(x)) dz$$

with respect to x_i and x_j for $1 \leq i < j \leq d$ we will have

$$\begin{aligned} & \frac{\partial(\mathcal{C}f)}{\partial x_i}(x, t) \\ &= \lambda(t) \int_0^1 \left(\left(1 + \frac{\partial \phi_i}{\partial x_i}(x, t, z)\right) \frac{\partial f}{\partial x_i}(x + \phi(x, t, z)) - \frac{\partial f}{\partial x_i}(x) \right. \\ & \quad - \frac{\partial \phi_1}{\partial x_i}(x, t, z) \frac{\partial f}{\partial x_1}(x) - \frac{\partial \phi_2}{\partial x_i}(x, t, z) \frac{\partial f}{\partial x_2}(x) - \cdots - \frac{\partial \phi_d}{\partial x_i}(x, t, z) \frac{\partial f}{\partial x_d}(x) \\ & \quad \left. - (\phi_1(x, t, z)) \frac{\partial^2 f}{\partial x_1 \partial x_i}(x) - (\phi_2(x, t, z)) \frac{\partial^2 f}{\partial x_2 \partial x_i}(x) - \cdots - (\phi_d(x, t, z)) \frac{\partial^2 f}{\partial x_d \partial x_i}(x) \right) dz \\ &= \lambda(t) \int_0^1 \left(\left(1 + \frac{\partial \phi_i}{\partial x_i}(x, t, z)\right) \frac{\partial f}{\partial x_i}(x + \phi(x, t, z)) - \frac{\partial f}{\partial x_i}(x) \right. \\ & \quad - \frac{\partial \phi_1}{\partial x_i}(x, t, z) \frac{\partial f}{\partial x_1}(x) - \frac{\partial \phi_2}{\partial x_i}(x, t, z) \frac{\partial f}{\partial x_2}(x) - \cdots - \frac{\partial \phi_d}{\partial x_i}(x, t, z) \frac{\partial f}{\partial x_d}(x) \\ & \quad \left. - (\phi_i(x, t, z)) \frac{\partial^2 f}{\partial x_i^2}(x) \right) dz \end{aligned}$$

by using the assumption $\frac{\partial^2 f}{\partial x_i \partial x_j}(x) = 0$ for $1 \leq i < j \leq d$. Then we will have

$$\begin{aligned} \frac{\partial^2(\mathcal{C}f)}{\partial x_i \partial x_j}(x, t) &= \lambda(t) \int_0^1 \left(\frac{\partial^2 \phi_i}{\partial x_i \partial x_j}(x, t, z) \frac{\partial f}{\partial x_i}(x + \phi(x, t, z)) \right. \\ & \quad + \left(1 + \frac{\partial \phi_i}{\partial x_i}(x, t, z)\right) \frac{\partial^2 f}{\partial x_i \partial x_j}(x + \phi(x, t, z)) \left(1 + \frac{\partial \phi_j}{\partial x_j}(x, t, z)\right) - \frac{\partial^2 f}{\partial x_i \partial x_j}(x) \\ & \quad - \frac{\partial^2 \phi_1}{\partial x_i \partial x_j}(x, t, z) \frac{\partial f}{\partial x_1}(x) - \cdots - \frac{\partial^2 \phi_d}{\partial x_i \partial x_j}(x, t, z) \frac{\partial f}{\partial x_d}(x) \\ & \quad \left. - \frac{\partial \phi_1}{\partial x_i}(x, t, z) \frac{\partial^2 f}{\partial x_1 \partial x_j}(x) - \cdots - \frac{\partial \phi_d}{\partial x_i}(x, t, z) \frac{\partial^2 f}{\partial x_d \partial x_j}(x) \right) dz \end{aligned}$$

$$\begin{aligned}
& -\frac{\partial \phi_i}{\partial x_j}(x, t, z) \frac{\partial^2 f}{\partial x_i^2}(x) - \phi_i(x, t, z) \frac{\partial^3 f}{\partial x_i^2 \partial x_j}(x) \Big) dz \\
& = \lambda(t) \int_0^1 \left(\frac{\partial^2 \phi_i}{\partial x_i \partial x_j}(x, t, z) \frac{\partial f}{\partial x_i}(x + \phi(x, t, z)) \right. \\
& \quad + (1 + \frac{\partial \phi_i}{\partial x_i}(x, t, z)) \frac{\partial^2 f}{\partial x_i \partial x_j}(x + \phi(x, t, z)) (1 + \frac{\partial \phi_j}{\partial x_j}(x, t, z)) \\
& \quad - \frac{\partial^2 \phi_1}{\partial x_i \partial x_j}(x, t, z) \frac{\partial f}{\partial x_1}(x) \cdots - \frac{\partial^2 \phi_d}{\partial x_i \partial x_j}(x, t, z) \frac{\partial f}{\partial x_d}(x) \\
& \quad \left. - \frac{\partial \phi_i}{\partial x_j}(x, t, z) \frac{\partial^2 f}{\partial x_i^2}(x) - \frac{\partial \phi_j}{\partial x_i}(x, t, z) \frac{\partial^2 f}{\partial x_j^2}(x) \right) dz \\
& = \lambda(t) \int_0^1 \left((1 + \frac{\partial \phi_i}{\partial x_i}(x, t, z)) (1 + \frac{\partial \phi_j}{\partial x_j}(x, t, z)) \frac{\partial^2 f}{\partial x_i \partial x_j}(x + \phi(x, t, z)) \right. \\
& \quad + \frac{\partial^2 \phi_i}{\partial x_i \partial x_j}(x, t, z) \frac{\partial f}{\partial x_i}(x + \phi(x, t, z)) - \sum_{k=1}^d \frac{\partial^2 \phi_k}{\partial x_i \partial x_j}(x, t, z) \frac{\partial f}{\partial x_k}(x) \\
& \quad \left. - \frac{\partial \phi_i}{\partial x_j}(x, t, z) \frac{\partial^2 f}{\partial x_i^2}(x) - \frac{\partial \phi_j}{\partial x_i}(x, t, z) \frac{\partial^2 f}{\partial x_j^2}(x) \right) dz
\end{aligned}$$

by using $\frac{\partial^2 f}{\partial x_i \partial x_j}(x) = 0$ for all $1 \leq i < j \leq d$.

□

Part II

Entropy and mutual information

Chapter 5

Information theory on the Wiener space

5.1 Introduction

The entropy of a d -dimensional Gaussian random vector can be calculated directly by using the definition of the differential entropy and the density function of a d -dimensional Gaussian random vector, and it is given in Proposition 1.5.9. In this chapter, first, we want to compare the mutual information generated by two Gaussian random vectors and obtain some equivalence relation. Second, we calculate the differential entropy of Brownian motion and its mutual information. Moreover, we derive a formula for the difference of differential entropies of multidimensional Brownian motion by using Itô calculus. Third, we state the definition of multidimensional bifractional Brownian motion, and then by using the Itô-Tanaka formula for d -dimensional bifractional Brownian motion we derive a formula for the difference of differential entropies.

This formula coincides with the formula that we derived for the difference of differential entropies of multidimensional Brownian motion with diffusion constants $K_i = 1$ and Hurst exponent $H_i = 1/2$, $1 \leq i \leq d$. In this Chapter the reason for these calculations is mainly theoretical: as we mostly recover known results by using Itô calculus. Finally, we end our chapter by providing bounds for Gaussian type densities.

5.2 Comparison of mutual information for Gaussian random vectors

In this section, we calculate the mutual information of two Gaussian random vectors. Then we derive necessary and sufficient conditions for comparison of mutual information of Gaussian random vectors, and after that we state necessary and sufficient conditions for 2-dimension as an example.

The differential entropy of a d -dimensional Gaussian random vector X with mean vector $\bar{\mu}$ and covariance matrix Σ is given in Proposition 1.5.9 as:

$$h(X_1, \dots, X_d) = \frac{1}{2} \log((2\pi e)^d \det \Sigma). \quad (5.2.1)$$

In the following lemma we calculate the mutual information of two Gaussian random vectors.

Lemma 5.2.1. *Let X and Y be d -dimensional Gaussian random vectors with covariance matrices Σ_{xx} and Σ_{yy} respectively and let (X, Y) be a pair*

CHAPTER 5. INFORMATION THEORY ON THE WIENER SPACE

of random vectors with covariance matrix Σ . Then we have

$$I(X, Y) = \frac{1}{2} \log \frac{\det(\Sigma_{xx}) \det(\Sigma_{yy})}{\det(\Sigma)}$$

$$\text{where } \Sigma = \begin{pmatrix} \Sigma_{xx} & \Sigma_{xy} \\ \Sigma_{yx} & \Sigma_{yy} \end{pmatrix} \text{ with } \Sigma_{xy} = \Sigma_{yx}^T.$$

Proof. From (5.2.1) we have

$$h(X) = \frac{1}{2} \log((2\pi e)^d \det(\Sigma_{xx}))$$

and

$$h(Y) = \frac{1}{2} \log((2\pi e)^d \det(\Sigma_{yy})).$$

Also we have

$$h(X, Y) = h(X_1, \dots, X_d, Y_1, \dots, Y_d) = \frac{1}{2} \log((2\pi e)^{2d} \det(\Sigma)).$$

Then we will have

$$\begin{aligned} I(X, Y) &= h(X) + h(Y) - h(X, Y) \\ &= \frac{1}{2} \log((2\pi e)^d \det(\Sigma_{xx})) + \frac{1}{2} \log((2\pi e)^d \det(\Sigma_{yy})) - \frac{1}{2} \log((2\pi e)^{2d} \det(\Sigma)) \\ &= \frac{1}{2} \log((2\pi e)^d \det(\Sigma_{xx})) + \frac{1}{2} \log((2\pi e)^d \det(\Sigma_{yy})) + \frac{1}{2} \log \frac{1}{(2\pi e)^{2d} \det(\Sigma)} \\ &= \frac{1}{2} \log \frac{\det(\Sigma_{xx}) \det(\Sigma_{yy})}{\det(\Sigma)}. \end{aligned}$$

□

In the following theorem necessary and sufficient conditions for com-

parison of mutual information of Gaussian random vectors is given.

Theorem 5.2.2. *Let X, X', Y , and Y' be d -dimensional Gaussian random vectors with covariance matrices $\Sigma_{xx}, \Sigma'_{xx}, \Sigma_{yy}$, and Σ'_{yy} respectively. Also let Σ and Σ' be covariance matrices of pairs (X, Y) and (X', Y') respectively. Then we have*

$$I(X, Y) \leq I(X', Y')$$

if and only if

$$\frac{\det(\Sigma_{xx}) \det(\Sigma_{yy})}{\det(\Sigma)} \leq \frac{\det(\Sigma'_{xx}) \det(\Sigma'_{yy})}{\det(\Sigma')}.$$

Proof. We have

$$\begin{aligned} I(X, Y) \leq I(X', Y') &\Leftrightarrow \frac{1}{2} \log \left(\frac{\det(\Sigma_{xx}) \det(\Sigma_{yy})}{\det(\Sigma)} \right) \leq \frac{1}{2} \log \left(\frac{\det(\Sigma'_{xx}) \det(\Sigma'_{yy})}{\det(\Sigma')} \right) \\ &\Leftrightarrow \frac{\det(\Sigma_{xx}) \det(\Sigma_{yy})}{\det(\Sigma)} \leq \frac{\det(\Sigma'_{xx}) \det(\Sigma'_{yy})}{\det(\Sigma')}. \end{aligned}$$

□

Example 5.2.1. Let $X = (X_1, X_2)$ and $Y = (Y_1, Y_2)$ be two Gaussian random vectors then we have the following comparison result:

$$I(X_1, X_2) \leq I(Y_1, Y_2) \Leftrightarrow \text{Corr}(X_1, X_2) \leq \text{Corr}(Y_1, Y_2)$$

where $\text{Corr}(X, Y)$ is the correlation between X and Y and it is defined by

$$\text{Corr}(X, Y) = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X)\text{Var}(Y)}}.$$

5.3 Differential entropy and mutual information of Brownian motion

In this section, first, we derive the differential entropy and mutual information of Brownian motion. Then we calculate the difference of differential entropies for multidimensional Brownian motion by using formula (5.2.1), and later we do the same derivation by using Itô calculus.

Consider a real valued Brownian motion $(B_t)_{t \in \mathbf{R}_+}$ and consider the vector $B = (B_{t_1}, \dots, B_{t_d})$ with its covariance matrix

$$\begin{aligned}
 & \begin{pmatrix} \text{Var}(B_{t_1}) & \text{Cov}(B_{t_1}, B_{t_2}) & \cdots & \text{Cov}(B_{t_1}, B_{t_d}) \\ \text{Cov}(B_{t_1}, B_{t_2}) & \text{Var}(B_{t_2}) & \cdots & \text{Cov}(B_{t_2}, B_{t_d}) \\ \vdots & \vdots & \cdots & \vdots \\ \text{Cov}(B_{t_1}, B_{t_d}) & \text{Cov}(B_{t_2}, B_{t_d}) & \cdots & \text{Var}(B_{t_d}) \end{pmatrix} \\
 &= \begin{pmatrix} t_1 & \min\{t_1, t_2\} & \cdots & \min\{t_1, t_d\} \\ \min\{t_1, t_2\} & t_2 & \cdots & \min\{t_2, t_d\} \\ \vdots & \vdots & \cdots & \vdots \\ \min\{t_1, t_d\} & \min\{t_2, t_d\} & \cdots & t_d \end{pmatrix} \\
 &= \begin{pmatrix} t_1 & t_1 & \cdots & t_1 \\ t_1 & t_2 & \cdots & t_2 \\ \vdots & \vdots & \cdots & \vdots \\ t_1 & t_2 & \cdots & t_d \end{pmatrix}
 \end{aligned}$$

for $0 < t_1 < \cdots < t_d$ then by Proposition 1.5.9 the differential entropy of

$B = (B_{t_1}, \dots, B_{t_d})$ is

$$h(B_{t_1}, \dots, B_{t_d}) = \frac{1}{2} \log \left((2\pi e)^d t_1 \prod_{i=2}^d (t_i - t_{i-1}) \right), \quad 0 < t_1 < t_2 < \dots < t_d. \quad (5.3.1)$$

We calculate the mutual information of $B = (B_{t_1}, \dots, B_{t_d})$ in the following lemma:

Lemma 5.3.1. *The mutual information of $B = (B_{t_1}, \dots, B_{t_d})$ is*

$$I(B_{t_1}, \dots, B_{t_d}) = -\frac{1}{2} \log \left(1 - \frac{t_{d-1}}{t_d} \right), \quad 0 < t_1 < t_2 < \dots < t_d.$$

Proof. We have

$$\begin{aligned} I(B_{t_1}, \dots, B_{t_d}) &= h(B_{t_1}, \dots, B_{t_{d-1}}) + h(B_{t_d}) - h(B_{t_1}, \dots, B_{t_d}) \\ &= \frac{1}{2} \log \left((2\pi e)^{d-1} t_1 \prod_{i=2}^{d-1} (t_i - t_{i-1}) \right) + \frac{1}{2} \log (2\pi e t_d) \\ &\quad - \frac{1}{2} \log \left((2\pi e)^d t_1 \prod_{i=2}^d (t_i - t_{i-1}) \right) \\ &= -\frac{1}{2} \log \left(1 - \frac{t_{d-1}}{t_d} \right). \end{aligned}$$

□

Differential entropies for multidimensional Brownian motion

Let $B_t = (B_t^1, \dots, B_t^d)$ be a d -dimensional standard Brownian motion in $t \in \mathbb{R}_+$ with covariance matrix

$$\Sigma_t = \begin{pmatrix} t & 0 & \cdots & 0 \\ 0 & t & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & t \end{pmatrix}.$$

Then from (5.2.1) we have

$$\begin{aligned} h(B_t^1, \dots, B_t^d) - h(B_s^1, \dots, B_s^d) &= \frac{1}{2} \log((2\pi e)^d \det \Sigma_t) - \frac{1}{2} \log((2\pi e)^d \det \Sigma_s) \\ &= \frac{1}{2} \log \frac{\det \Sigma_t}{\det \Sigma_s} \\ &= \frac{1}{2} \log \frac{t^d}{s^d} \\ &= \frac{d}{2} \log \frac{t}{s}. \end{aligned}$$

In the following proposition we calculate the difference of differential entropies for multidimensional Brownian motion by using Itô calculus, an alternative proof.

Proposition 5.3.2. *Let $B = (B_t^1, \dots, B_t^d)$ be a d -dimensional Brownian motion with covariance matrix Σ . Then we have*

$$h(B_t^1, \dots, B_t^d) - h(B_s^1, \dots, B_s^d) = \frac{d}{2} \log \frac{t}{s}, \quad 0 < s \leq t.$$

Proof. Denote $\Sigma = (\alpha_{ij})$ and its inverse $\Sigma^{-1} = (\beta_{ij})$. The probability

density function of $(B_u^1, \dots, B_u^d) \sim N(0, u\Sigma)$ is

$$p_u(x) = \frac{1}{(2\pi)^{d/2}(\det(u\Sigma))^{1/2}} \exp\left(-\frac{x^T(u\Sigma)^{-1}x}{2}\right). \quad (5.3.2)$$

By taking logarithm in both sides of (5.3.2) we will have

$$\log p_u(x) = -\frac{d}{2} \log 2\pi - \frac{1}{2} \log(u \det \Sigma) - \frac{x^T \Sigma^{-1} x}{2u}. \quad (5.3.3)$$

After taking partial derivatives in both sides of (5.3.3) we will have

$$\frac{\partial^2 \log p_u}{\partial x_i \partial x_j}(x) = -\frac{\beta_{ij}}{u}. \quad (5.3.4)$$

We use the multidimensional Itô formula

$$\begin{aligned} & d \log p_u(B_u^1, \dots, B_u^d) \\ &= \frac{\partial \log p_u}{\partial u}(B_u^1, \dots, B_u^d) du \\ &+ \sum_{i=1}^d \frac{\partial \log p_u}{\partial x_i}(B_u^1, \dots, B_u^d) dB_u^i + \frac{1}{2} \sum_{i=1}^d \sum_{j=1}^d \frac{\partial^2 \log p_u}{\partial x_i \partial x_j}(B_u^1, \dots, B_u^d) d\langle B^i, B^j \rangle_u \\ &= \frac{\partial \log p_u}{\partial u}(B_u^1, \dots, B_u^d) du \\ &+ \sum_{i=1}^d \frac{\partial \log p_u}{\partial x_i}(B_u^1, \dots, B_u^d) dB_u^i + \frac{1}{2} \sum_{i=1}^d \sum_{j=1}^d \frac{\partial^2 \log p_u}{\partial x_i \partial x_j}(B_u^1, \dots, B_u^d) \alpha_{ij} du. \end{aligned}$$

Then we will have

$$\begin{aligned} \log p_t(B_t^1, \dots, B_t^d) - \log p_s(B_s^1, \dots, B_s^d) &= \int_s^t \left(\frac{\partial \log p_u}{\partial u}(B_u^1, \dots, B_u^d) \right) du \\ &+ \int_s^t \sum_{i=1}^d \left(\frac{\partial \log p_u}{\partial x_i}(B_u^1, \dots, B_u^d) \right) dB_u^i \end{aligned}$$

$$+ \int_s^t \frac{1}{2} \sum_{i=1}^d \sum_{j=1}^d \left(\frac{\partial^2 \log p_u}{\partial x_i \partial x_j} (B_u^1, \dots, B_u^d) \right) \alpha_{ij} du. \quad (5.3.5)$$

After taking expectations in both sides of (5.3.5) and by using (5.3.4) we will have

$$\begin{aligned} & h(B_t^1, \dots, B_t^d) - h(B_s^1, \dots, B_s^d) \\ &= -\mathbb{E}[\log p_t(B_t^1, \dots, B_t^d) - \log p_s(B_s^1, \dots, B_s^d)] \\ &= -\mathbb{E} \left[\int_s^t \left(\frac{\partial \log p_u}{\partial u} (B_u^1, \dots, B_u^d) \right) du + \int_s^t \frac{1}{2} \sum_{i=1}^d \sum_{j=1}^d \left(\frac{\partial^2 \log p_u}{\partial x_i \partial x_j} (B_u^1, \dots, B_u^d) \right) \alpha_{ij} du \right] \\ &= -\mathbb{E} \left[\int_s^t \frac{1}{2} \sum_{i=1}^d \sum_{j=1}^d \left(\frac{\partial^2 \log p_u}{\partial x_i \partial x_j} (B_u^1, \dots, B_u^d) \right) \alpha_{ij} du \right] \\ &= \mathbb{E} \left[\int_s^t \frac{1}{2u} \sum_{i=1}^d \sum_{j=1}^d \beta_{ij} \alpha_{ij} du \right] = \frac{d}{2} \log \frac{t}{s}, \end{aligned}$$

since

$$\begin{aligned} \mathbb{E} \left[\int_s^t \frac{\partial \log p_u}{\partial u} (B_u^1, \dots, B_u^d) du \right] &= \mathbb{E} \left[\int_s^t \frac{1}{p_u(x)} \frac{\partial p_u}{\partial u} (B_u^1, \dots, B_u^d) du \right] \\ &= \int_{\mathbb{R}^d} \left(\int_s^t \frac{1}{p_u(x)} \frac{\partial p_u}{\partial u} (x) du \right) p_u(x) dx \\ &= \int_s^t \int_{\mathbb{R}^d} \left(\frac{1}{p_u(x)} \frac{\partial p_u}{\partial u} (x) \right) p_u(x) dx du \\ &= \int_s^t \frac{\partial}{\partial u} \int_{\mathbb{R}^d} p_u(x) dx du \\ &= 0. \end{aligned}$$

□

5.4 Differential entropies for multidimensional bifractional Brownian motion

In this section, we calculate difference of differential entropies for multidimensional bifractional Brownian motion. First we define multidimensional bifractional Brownian motion, and then by using the Itô-Tanaka formula for multidimensional bifractional Brownian motion we calculate difference of differential entropies. We follow notations of Es-Sebaïy and Tudor [12].

The bifractional Brownian motion $B^{H,K}$ is a centered Gaussian process, starting from zero, with covariance function

$$R^{H,K}(t, s) := R(t, s) = \frac{1}{2^K} ((t^{2H} + s^{2H})^K - |t - s|^{2HK})$$

where the parameters H and K are such that $H \in (0, 1)$ and $K \in (0, 1]$. In the case $K = 1$ we retrieve the fractional Brownian motion while the case $K = 1$ and $H = 1/2$ corresponds to the standard Brownian motion.

Let $(B^{H,K}, t \in [0, T])$ be a bifractional Brownian motion on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and let $\mathcal{H}$ be a Hilbert space associated to the bifractional Brownian motion. This space is defined as the completion of the linear space $\mathcal{E}$ generated by the indicator functions $1_{[0,t]}, t \in [0, T]$ with respect to the inner product

$$\langle 1_{[0,t]}, 1_{[0,s]} \rangle_{\mathcal{H}} = R(t, s).$$

The application $\psi \in \mathcal{E} \rightarrow B^{H,K}(\psi)$ is an isometry from $\mathcal{E}$ to the Gaussian

space generated by $B^{H,K}$ and it can be extended to $\mathcal{H}$.

Let us denote by $\mathcal{S}$ the set of smooth functionals of the form

$$F = f(B^{H,K}(\psi_1), \dots, B^{H,K}(\psi_n)),$$

where $f \in C_b^\infty(\mathbb{R}^n)$ and $\psi_i \in \mathcal{H}$. The Malliavin derivative of the functional F as above is given by

$$D^{B^{H,K}} F = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(B^{H,K}(\psi_1), \dots, B^{H,K}(\psi_n)) \psi_i.$$

The divergence (or Skorohod) integral $\delta^{B^{H,K}}$ is the adjoint operator of $D^{B^{H,K}}$. Concretely, a random variable $u \in L^2(\Omega, \mathcal{H})$ belongs to the domain of the divergence operator $Dom(\delta^{B^{H,K}})$ if

$$\mathbb{E}[|\langle D^{B^{H,K}} F, u \rangle_{\mathcal{H}}|] \leq c \|F\|_{L^2(\Omega)}$$

for every $F \in \mathcal{S}$. In this case $\delta^{B^{H,K}}(u)$ is given by the duality relationship

$$\mathbb{E}[F \delta^{B^{H,K}}(u)] = \mathbb{E}[\langle D^{B^{H,K}} F, u \rangle_{\mathcal{H}}]$$

for any $F \in \mathbb{D}^{1,2}$. Here $\mathbb{D}^{m,2}(m \geq 1)$ is the closure of $\mathcal{S}$.

Note that the Skorohod integral is a continuous operator from $\mathbb{D}^{\alpha,p}(\mathcal{H})$ into $\mathbb{D}^{\alpha,p}$ for $p > 1$ and $\alpha \in \mathbb{R}$. For adapted integrands, the divergence integral coincides with the classical Itô integral. We will use the notation

$$\delta^{B^{H,K}}(u) = \int_0^T u_s \delta B_s^{H,K}.$$

The Itô-Tanaka formula is given in Theorem 4 of [12] in the following way:

Let $f \in C^{1,2}([0, T] \times \mathbb{R}^d, \mathbb{R})$ and $B_t^{H,K} = (B_t^{H_1, K_1}, \dots, B_t^{H_d, K_d})$ be a d -dimensional bifractional Brownian motion with $2H_i K_i \geq 1$ for any $i = 1, \dots, d$. Then for every i one has $\frac{\partial f}{\partial x_i}(\cdot, B^{H,K}) \in \text{Dom}(\delta^{B^{H_i, K_i}})$ and for every t

$$\begin{aligned} f(t, B_t^{H,K}) = & f(0, 0) + \int_0^t \frac{\partial f}{\partial u}(u, B_u^{H,K}) du + \sum_{i=1}^d \int_0^t \frac{\partial f}{\partial x_i}(u, B_u^{H,K}) \delta B_u^{H_i, K_i} \\ & + \left(\sum_{i=1}^d H_i K_i \right) \int_0^t \frac{\partial^2 f}{\partial x_i^2}(u, B_u^{H,K}) u^{2H_i K_i - 1} du. \end{aligned} \quad (5.4.1)$$

Given two vectors $H = (H_1, \dots, H_d) \in (0, 1)^d$ and $K = (K_1, \dots, K_d) \in (0, 1]^d$ we introduce the d -dimensional bifractional Brownian motion

$$B^{H,K} = (B^{H_1, K_1}, \dots, B^{H_d, K_d})$$

as a centered Gaussian vector whose components are independent one-dimensional bifractional Brownian motions.

Then we have the following proposition for difference of differential entropies of multidimensional bifractional Brownian motion:

Proposition 5.4.1. *Let $(B_t^{H,K})_{t \in \mathbb{R}_+}$ be a d -dimensional bifractional Brownian motion and let $2H_i K_i \geq 1$. Then we have*

$$h(B_t^{H_1, K_1}, \dots, B_t^{H_d, K_d}) - h(B_s^{H_1, K_1}, \dots, B_s^{H_d, K_d}) = \left(\sum_{i=1}^d H_i K_i \right) \log \frac{t}{s}, \quad 0 < s \leq t. \quad (5.4.2)$$

Proof. The density function of the multidimensional bifractional Brown-

ian motion with covariance matrix Σ is

$$p_u(x) = \frac{1}{(2\pi)^{d/2}(\det \Sigma)^{1/2}} \exp\left(-\frac{x^T \Sigma^{-1} x}{2}\right)$$

$$\text{where } \Sigma = \begin{pmatrix} u^{2H_1 K_1} & 0 & \dots & 0 \\ 0 & u^{2H_2 K_2} & \dots & 0 \\ \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & u^{2H_d K_d} \end{pmatrix}.$$

After taking the logarithm of the density function and differentiating it with respect to x_i twice we will have

$$\frac{\partial^2 \log p_u}{\partial x_i^2}(x) = -u^{-2H_i K_i}, \quad 1 \leq i \leq d.$$

Then by using the Itô-Tanaka formula (5.4.1) we will have

$$\begin{aligned} & h(B_t^{H_1, K_1}, \dots, B_t^{H_d, K_d}) - h(B_s^{H_1, K_1}, \dots, B_s^{H_d, K_d}) \\ &= -\mathbb{E} \left[\sum_{i=1}^d H_i K_i \int_s^t \frac{\partial^2 \log p_u}{\partial x_i^2}(B_u^{H_1, K_1}, \dots, B_u^{H_d, K_d}) u^{2H_i K_i - 1} du \right] \\ &= \mathbb{E} \left[\sum_{i=1}^d H_i K_i \int_s^t u^{-2H_i K_i} u^{2H_i K_i - 1} du \right] \\ &= \left(\sum_{i=1}^d H_i K_i \right) \int_s^t \frac{1}{u} du \\ &= \left(\sum_{i=1}^d H_i K_i \right) \log \frac{t}{s}, \quad 0 < s \leq t. \end{aligned}$$

□

Remark 5.4.2. We can derive the result in Proposition 5.4.1 by using

(5.2.1).

Remark 5.4.3. *We can derive the result in Proposition 5.3.2 by taking $H_i = 1/2$ and $K_i = 1$ for all $1 \leq i \leq d$ in Proposition 5.4.1.*

5.5 Log-Sobolev type inequalities for Gaussian densities

In this section we derive bounds for differential entropies of Gaussian densities.

Let F be a ‘smooth random variable’ of the form

$$F = f \left(\int_0^\infty u_1(t) dB_t, \dots, \int_0^\infty u_d(t) dB_t \right) \quad (5.5.1)$$

where $f \in C_b^1(\mathbb{R}^d)$ and $u_1, \dots, u_d \in L^2(\mathbb{R}_+)$ and $(B_t)_{t \in \mathbb{R}_+}$ is a standard Brownian motion. Then the Malliavin gradient of a smooth random variable F of the form (5.5.1) is given by

$$D_t F = \sum_{i=1}^d \partial_i f \left(\int_0^\infty u_1(t) dB_t, \dots, \int_0^\infty u_d(t) dB_t \right) u_i(t), \quad t \geq 0. \quad (5.5.2)$$

For more properties of the operator D and its relation with the Clark-Ocone formula one can check Nguyen et al. [28].

A general expression for Gaussian type densities is given in (1.3) of [29] as follows

$$P_z(F) = \frac{\mathbb{E}[|F|]}{2\varphi_F(z)} \exp \left(- \int_0^z \frac{u}{\varphi_F(u)} du \right) \quad (5.5.3)$$

where

$$\varphi_F(z) = \mathbb{E} \left[\int_0^\infty D_t F \mathbb{E}[D_t F | \mathcal{F}_t] \Big| F = z \right]. \quad (5.5.4)$$

If we take $z = F$ in (5.5.4) we will have

$$\varphi_F(F) = \mathbb{E} \left[\int_0^\infty D_t F \mathbb{E}[D_t F | \mathcal{F}_t] \Big| F \right].$$

In the following propositions we derive lower and upper bounds for the differential entropy of Gaussian type densities.

Proposition 5.5.1. *Assume that we have two deterministic functions $f_1(t)$ and $f_2(t)$ such that*

$$f_1(t) \leq D_t F \leq f_2(t)$$

where F is defined in (5.5.1) with its Malliavin derivative in (5.5.2).

Then we will have

$$\frac{\mathbb{E}[F^2]}{2f_2(t)^2} + \log f_1(t)^2 - \log \frac{\mathbb{E}[|F|]}{2} \leq h(F) \leq \frac{\mathbb{E}[F^2]}{2f_1(t)^2} + \log f_2(t)^2 - \log \frac{\mathbb{E}[|F|]}{2}.$$

Proof. Under the assumption

$$f_1(t) \leq D_t F \leq f_2(t)$$

if we use (5.5.4) we will have

$$f_1(t)^2 \leq \mathbb{E} \left[\int_0^\infty D_t F \mathbb{E}[D_t F | \mathcal{F}_t] \Big| F \right] \leq f_2(t)^2$$

i.e.

$$f_1(t)^2 \leq \varphi_F(F) \leq f_2(t)^2.$$

Then by using (5.5.3) we will have

$$\frac{\mathbb{E}[|F|]}{2f_2(t)^2} \exp\left(-\frac{z^2}{2f_1(t)^2}\right) \leq P_z(F) \leq \frac{\mathbb{E}[|F|]}{2f_1(t)^2} \exp\left(-\frac{z^2}{2f_2(t)^2}\right). \quad (5.5.5)$$

After taking logarithm in each side of (5.5.5) we will have

$$\log \frac{\mathbb{E}[|F|]}{2} - \log f_2(t)^2 - \frac{z^2}{2f_1(t)^2} \leq \log P_z(F) \leq \log \frac{\mathbb{E}[|F|]}{2} - \log f_1(t)^2 - \frac{z^2}{2f_2(t)^2}$$

then we have

$$\frac{z^2}{2f_2(t)^2} + \log f_1(t)^2 - \log \frac{\mathbb{E}[|F|]}{2} \leq -\log P_z(F) \leq \frac{z^2}{2f_1(t)^2} + \log f_2(t)^2 - \log \frac{\mathbb{E}[|F|]}{2}. \quad (5.5.6)$$

By taking expectation in each side of (5.5.6) we will have

$$\frac{\mathbb{E}[F^2]}{2f_2(t)^2} + \log f_1(t)^2 - \log \frac{\mathbb{E}[|F|]}{2} \leq h(F) \leq \frac{\mathbb{E}[F^2]}{2f_1(t)^2} + \log f_2(t)^2 - \log \frac{\mathbb{E}[|F|]}{2}.$$

□

Proposition 5.5.2. *Assume that $\sigma_{\min}$ and $\sigma_{\max}$ are two nonnegative constants with*

$$\sigma_{\min}^2 \leq \varphi_F(F) \leq \sigma_{\max}^2 \quad (5.5.7)$$

where F is defined in (5.5.1) with its Malliavin derivative in (5.5.2).

Then we have

$$\frac{\mathbb{E}[F^2]}{2\sigma_{\max}^2} + 2\log \sigma_{\min} - \log \frac{\mathbb{E}[|F|]}{2} \leq h(F) \leq \frac{\mathbb{E}[F^2]}{2\sigma_{\min}^2} + 2\log \sigma_{\max} - \log \frac{\mathbb{E}[|F|]}{2}. \quad (5.5.8)$$

Proof. By using (5.5.3) under the restriction (5.5.7) we will have

$$\frac{\mathbb{E}[|F|]}{2\sigma_{\max}^2} \exp\left(-\frac{z^2}{2\sigma_{\min}^2}\right) \leq P_z(F) \leq \frac{\mathbb{E}[|F|]}{2\sigma_{\min}^2} \exp\left(-\frac{z^2}{2\sigma_{\max}^2}\right). \quad (5.5.9)$$

Then by taking logarithm in each side of (5.5.9) we will have

$$\log \frac{\mathbb{E}[|F|]}{2} - 2\log \sigma_{\max} - \frac{z^2}{2\sigma_{\min}^2} \leq \log P_z(F) \leq \log \frac{\mathbb{E}[|F|]}{2} - 2\log \sigma_{\min} - \frac{z^2}{2\sigma_{\max}^2}$$

then we have

$$\frac{z^2}{2\sigma_{\max}^2} + 2\log \sigma_{\min} - \log \frac{\mathbb{E}[|F|]}{2} \leq -\log P_z(F) \leq \frac{z^2}{2\sigma_{\min}^2} + 2\log \sigma_{\max} - \log \frac{\mathbb{E}[|F|]}{2}. \quad (5.5.10)$$

After taking expectation in each side of (5.5.10) we will have

$$\frac{\mathbb{E}[F^2]}{2\sigma_{\max}^2} + 2\log \sigma_{\min} - \log \frac{\mathbb{E}[|F|]}{2} \leq h(F) \leq \frac{\mathbb{E}[F^2]}{2\sigma_{\min}^2} + 2\log \sigma_{\max} - \log \frac{\mathbb{E}[|F|]}{2}.$$

□

In the following example we take $F = B_T \sim N(0, T)$ and we find that lower and upper bounds coincide.

Example 5.5.3. Let $F = B_T$ then we have $D_t B_T = 1_{[0, T]}(t)$ and $\varphi_F(F) =$

T. Since $\mathbb{E}[|F|] = \frac{\sqrt{2T}}{\sqrt{\pi}}$ we will have

$$h(F) = h(B_T) = \frac{T}{2T} + \log T - \log \frac{\sqrt{2T}}{2\sqrt{\pi}} = \frac{1}{2} \log(2e\pi T).$$

Chapter 6

Information theory on the Poisson space

6.1 Introduction

In this chapter, first, we state the entropy of a Poisson random variable and then we derive a formula between the information entropy and the probability entropy for Poisson random variables. Moreover, in Section 6.3, we obtain a similar formula which connects the information entropy and the probability entropy for Poisson process; we use Itô calculus. The reason for these calculations is purely theoretical; first we obtained a relation between Shannon entropy and probability entropy and then we use Itô calculus to verify it. Second, we calculate the entropy of the compound Poisson process and the Gamma process. Third, we bound the relative entropy of two probability measures on the Poisson space, and then we derive the logarithmic Sobolev inequality for Poisson type densities. Finally, we provide calculations of predictable decompositions.

6.2 Entropy of Poisson random variables

In this section we derive a formula between the Shannon entropy and the probability entropy.

The entropy of a Poisson random variable X_λ with parameter λ is given by

$$H(X_\lambda) = - \sum_x \mathbb{P}(X_\lambda = x) \log \mathbb{P}(X_\lambda = x) = \lambda - \lambda \log \lambda + \sum_{k=0}^{\infty} \frac{e^{-\lambda} \lambda^k}{k!} \log k!. \quad (6.2.1)$$

Recall that the probability entropy of a non-negative random variable is defined by

$$\text{Ent}[X_\lambda] = \mathbb{E}[X_\lambda \log X_\lambda] - \mathbb{E}[X_\lambda] \log \mathbb{E}[X_\lambda]. \quad (6.2.2)$$

One can show the nonnegativity of the probability entropy by applying Jensens' inequality to the convex function $\Phi(x) = x \log x$.

We have the following connection formula between the information entropy and the probability entropy for Poisson random variables:

Proposition 6.2.1. *Let $X_\lambda \sim P(\lambda)$ then we have*

$$\frac{\partial H(X_\lambda)}{\partial \lambda} = \frac{1}{\lambda} \text{Ent}[X_\lambda]. \quad (6.2.3)$$

Proof. We have

$$\frac{\partial H(X_\lambda)}{\partial \lambda} = \frac{\partial}{\partial \lambda} \left(\lambda - \lambda \log \lambda + \sum_{k=0}^{\infty} \frac{e^{-\lambda} \lambda^k}{k!} \log k! \right)$$

$$\begin{aligned}
&= -\log \lambda + \sum_{x=0}^{\infty} \frac{e^{-\lambda} \lambda^x}{x!} \log(x+1) \\
&= -\log \lambda + \sum_{x=1}^{\infty} \frac{e^{-\lambda} \lambda^{x-1}}{(x-1)!} \log x \\
&= -\log \lambda + \lambda^{-1} \sum_{x=1}^{\infty} \frac{e^{-\lambda} \lambda^x}{(x-1)!} \log x \\
&= -\log \lambda + \lambda^{-1} \sum_{x=0}^{\infty} \frac{e^{-\lambda} \lambda^x}{x!} x \log x \\
&= -\log \lambda + \lambda^{-1} \mathbb{E}[X_\lambda \log X_\lambda] \\
&= -\log \lambda + \lambda^{-1} (\mathbb{E}[X_\lambda \log X_\lambda] - \mathbb{E}[X_\lambda] \log \mathbb{E}[X_\lambda] + \mathbb{E}[X_\lambda] \log \mathbb{E}[X_\lambda]) \\
&= -\log \lambda + \lambda^{-1} \mathbb{E}[X_\lambda] \log \mathbb{E}[X_\lambda] + \lambda^{-1} (\mathbb{E}[X_\lambda \log X_\lambda] - \mathbb{E}[X_\lambda] \log \mathbb{E}[X_\lambda]) \\
&= -\log \lambda + \lambda^{-1} \mathbb{E}[X_\lambda] \log \mathbb{E}[X_\lambda] + \lambda^{-1} \text{Ent}[X_\lambda] \\
&= \frac{1}{\lambda} \text{Ent}[X_\lambda],
\end{aligned}$$

by using (6.2.1) and (6.2.2). □

Let $(X_\lambda)_{\lambda \in [\lambda_1, \lambda_2]}$ be a collection of Poisson random variables then we have the following formula for difference of entropies:

Proposition 6.2.2. *Let $X_{\lambda_1} \sim P(\lambda_1)$ and $X_{\lambda_2} \sim P(\lambda_2)$ with $0 < \lambda_1 \leq \lambda_2$. Then we have*

$$H(X_{\lambda_2}) - H(X_{\lambda_1}) = \int_{\lambda_1}^{\lambda_2} \frac{1}{u} \text{Ent}[X_u] du.$$

Proof. Let $X_u \sim P(u)$ with $\lambda_1 \leq u \leq \lambda_2$. From (6.2.3) we have

$$\frac{\partial H(X_u)}{\partial u} = \frac{1}{u} \text{Ent}[X_u].$$

Then

$$\int_{\lambda_1}^{\lambda_2} \frac{\partial H(X_u)}{\partial u} du = \int_{\lambda_1}^{\lambda_2} \frac{1}{u} \text{Ent}[X_u] du$$

implies

$$H(X_{\lambda_2}) - H(X_{\lambda_1}) = \int_{\lambda_1}^{\lambda_2} \frac{1}{u} \text{Ent}[X_u] du.$$

□

6.3 Entropy of Poisson process

In the following proposition we recover a result similar to the result in Proposition 6.2.1 for Poisson process by using Itô calculus.

Proposition 6.3.1. *Let $N_t \sim P(\lambda t)$ then we have*

$$\frac{\partial H(N_t)}{\partial t} = \frac{1}{t} \text{Ent}[N_t].$$

Proof. Denote our probability mass function by $p_t(k)$:

$$p_t(k) = e^{-\lambda t} \frac{(\lambda t)^k}{k!}, \quad k = 0, 1, 2, \dots$$

By using the Itô formula we have

$$\begin{aligned} \log p_t(N_t) &= \log p_0(N_0) \\ &+ \int_0^t (\log p_s(N_s + 1) - \log p_s(N_s)) (dN_s - \lambda ds) \\ &+ \lambda \int_0^t (\log p_s(N_s + 1) - \log p_s(N_s)) ds \\ &+ \int_0^t \frac{\partial \log p_s}{\partial s}(N_s) ds. \end{aligned}$$

Then we have

$$\begin{aligned} H(N_t) &= -\mathbb{E}[\log p_t(N_t)] \\ &= -\mathbb{E} \left[\lambda \int_0^t (\log p_s(N_s + 1) - \log p_s(N_s)) ds + \int_0^t \frac{1}{p_s(N_s)} \frac{\partial p_s}{\partial s}(N_s) ds \right]. \end{aligned} \quad (6.3.1)$$

Taking derivative in both sides of (6.3.1) with respect to t we will have

$$\frac{\partial H(N_t)}{\partial t} = -\mathbb{E} \left[\lambda (\log p_t(N_t + 1) - \log p_t(N_t)) + \frac{1}{p_t(N_t)} \frac{\partial p_t}{\partial t}(N_t) \right].$$

Since $p_t(k) = e^{-\lambda t} \frac{(\lambda t)^k}{k!}$ then we can write

$$p_t(N_t + 1) = e^{-\lambda t} \frac{(\lambda t)^{N_t+1}}{(N_t + 1)!} \text{ and } p_t(N_t) = e^{-\lambda t} \frac{(\lambda t)^{N_t}}{(N_t)!}.$$

Then we have

$$\begin{aligned} &\log p_t(N_t + 1) - \log p_t(N_t) \\ &= (-\lambda t + (N_t + 1) \log \lambda t - \log(N_t + 1)!) - (-\lambda t + N_t \log \lambda t - \log(N_t)!) \\ &= \log \lambda t - \log(N_t + 1). \end{aligned}$$

Also we have

$$\begin{aligned} \mathbb{E} \left[\frac{1}{p_t(N_t)} \frac{\partial p_t}{\partial t}(N_t) \right] &= \sum_x \left(\frac{1}{p_t(x)} \frac{\partial p_t}{\partial t}(x) \right) p_t(x) \\ &= \sum_x \frac{\partial p_t}{\partial t}(x) \\ &= \frac{\partial}{\partial t} \sum_x p_t(x) \end{aligned}$$

$$= 0.$$

Finally we have

$$\begin{aligned} \frac{\partial H(N_t)}{\partial t} &= -\mathbb{E}[\lambda(\log \lambda t - \log(N_t + 1))] \\ &= -\lambda \log \lambda t + \lambda \mathbb{E}[\log(N_t + 1)] \\ &= -\lambda \log \lambda t + \lambda \sum_{x=0}^{\infty} \frac{e^{-\lambda t} (\lambda t)^x}{x!} \log(x + 1) \\ &= -\lambda \log \lambda t + \frac{1}{t} \sum_{x=1}^{\infty} \frac{e^{-\lambda t} (\lambda t)^x}{x!} x \log x \\ &= -\lambda \log \lambda t + \frac{1}{t} \mathbb{E}[N_t \log N_t] \\ &= -\lambda \log \lambda t + \frac{1}{t} \mathbb{E}[N_t] \log \mathbb{E}[N_t] + \frac{1}{t} \text{Ent}[N_t] \\ &= \frac{1}{t} \text{Ent}[N_t]. \end{aligned}$$

□

Remark 6.3.2. *Proposition 6.3.1 is a special case of (6.2.3) with the random variable $N_t \sim P(\lambda t)$.*

6.4 Entropy of the compound Poisson and the Gamma processes

In this section, we define the compound Poisson process and its distribution and then we calculate the entropy of the compound Poisson process and the Gamma process.

Entropy of the compound Poisson process

Let $(Z_k)_{k \geq 1}$ denote an i.i.d. sequence of square-integrable random variables with probability distribution $\nu(dy)$ on $\mathbb{R}$, independent of the Poisson process $(N_t)_{t \in \mathbb{R}_+}$. Then we have

$$\mathbb{P}(Z_k \in [a, b]) = \nu([a, b]) = \int_a^b \nu(dy), \quad -\infty < a \leq b < \infty.$$

The following sum

$$Y_t = \sum_{k=1}^{N_t} Z_k, \quad t \in \mathbb{R}_+ \tag{6.4.1}$$

is called the compound Poisson process with distribution function

$$P(Y_t \in A) = e^{-t\nu(\mathbb{R})} \sum_{k=0}^{\infty} \frac{t^k \nu^k}{k!}(A) \tag{6.4.2}$$

where ν is the Lévy measure of Y_t and ν^k is the k -fold convolution of ν with $\nu(\mathbb{R}) = 1$. For more details on the compound Poisson process one can check the book of Sato [34].

The entropy relation for the compound Poisson process is stated in the following lemma.

Lemma 6.4.1. *Let Y_t be a compound Poisson process as defined in (6.4.1).*

Then we have

$$\frac{\partial H(Y_t)}{\partial t} = -\mathbb{E} \left[\lambda \int_{-\infty}^{\infty} (\log P_t(Y_t + x) - \log P_t(Y_t)) \nu(dx) \right].$$

Proof. We will use Itô formula for jump processes (see Chapter 4 of Ap-

plebaum [1])

$$\begin{aligned}
 \log P_t(Y_t) &= \log P_0(Y_0) \\
 &+ \int_0^t \int_{-\infty}^{\infty} (\log P_s(Y_s + x) - \log P_s(Y_s)) (w(dx, ds) - \lambda \nu(dx) ds) \\
 &+ \lambda \int_0^t \int_{-\infty}^{\infty} (\log P_s(Y_s + x) - \log P_s(Y_s)) \nu(dx) ds \\
 &+ \int_0^t \frac{\partial \log P_s}{\partial s}(Y_s) ds.
 \end{aligned}$$

Then we have

$$\begin{aligned}
 H(Y_t) &= -\mathbb{E} [\log P_t(Y_t)] \\
 &= -\mathbb{E} \left[\lambda \int_0^t \int_{-\infty}^{\infty} (\log P_s(Y_s + x) - \log P_s(Y_s)) \nu(dx) ds + \int_0^t \frac{\partial \log P_s}{\partial s}(Y_s) ds \right].
 \end{aligned} \tag{6.4.3}$$

By taking derivative in both sides of (6.4.3) with respect to t we will have

$$\begin{aligned}
 \frac{\partial H(Y_t)}{\partial t} &= -\mathbb{E} \left[\lambda \int_{-\infty}^{\infty} (\log P_t(Y_t + x) - \log P_t(Y_t)) \nu(dx) + \frac{1}{P_t(Y_t)} \frac{\partial P_t}{\partial t}(Y_t) \right] \\
 &= -\mathbb{E} \left[\lambda \int_{-\infty}^{\infty} (\log P_t(Y_t + x) - \log P_t(Y_t)) \nu(dx) \right] - \mathbb{E} \left[\frac{1}{P_t(Y_t)} \frac{\partial P_t}{\partial t}(Y_t) \right] \\
 &= -\mathbb{E} \left[\lambda \int_{-\infty}^{\infty} (\log P_t(Y_t + x) - \log P_t(Y_t)) \nu(dx) \right]
 \end{aligned}$$

since

$$\begin{aligned}
 \mathbb{E} \left[\frac{1}{P_t(Y_t)} \frac{\partial P_t}{\partial t}(Y_t) \right] &= \sum_x \left(\frac{1}{P_t(x)} \frac{\partial P_t}{\partial t}(x) \right) P_t(x) \\
 &= \sum_x \frac{\partial P_t}{\partial t}(x) \\
 &= \frac{\partial}{\partial t} \sum_x P_t(x)
 \end{aligned}$$

$$= 0.$$

□

Entropy of the Gamma process

The Gamma process is a stochastic process with independent and stationary increments which has the gamma distribution with the density function:

$$x \mapsto \frac{\lambda^t x^{t-1} e^{-\lambda x}}{\Gamma(t)}. \quad (6.4.4)$$

In the following lemma we state the entropy relation of the Gamma process.

Lemma 6.4.2. *The entropy relation of the Gamma process is as follows*

$$\frac{\partial H(Y_t)}{\partial t} = \lambda + \lambda(1-t) \mathbb{E} \left[\int_0^\infty (\log(Y_t + x) - \log Y_t) \frac{e^{-\lambda x}}{x} dx \right]. \quad (6.4.5)$$

Proof. For the Gamma process we have

$$P_t(x) = \frac{\lambda^t x^{t-1} e^{-\lambda x}}{\Gamma(t)} \text{ and } \nu(dx) = \frac{e^{-\lambda x}}{x} 1_{\{x>0\}} dx.$$

Then we can write

$$P_t(Y_t) = \frac{\lambda^t Y_t^{t-1} e^{-\lambda Y_t}}{\Gamma(t)} \text{ and } P_t(Y_t + x) = \frac{\lambda^t (Y_t + x)^{t-1} e^{-\lambda(Y_t+x)}}{\Gamma(t)},$$

which implies

$$\begin{aligned}\log P_t(Y_t) &= t \log \lambda + (t-1) \log Y_t - \lambda Y_t - \log \Gamma(t) \\ \log P_t(Y_t + x) &= t \log \lambda + (t-1) \log(Y_t + x) - \lambda(Y_t + x) - \log \Gamma(t).\end{aligned}$$

Then

$$\log P_t(Y_t + x) - \log P_t(Y_t) = (t-1)(\log(Y_t + x) - \log Y_t) - \lambda x.$$

Finally we will have

$$\begin{aligned}\frac{\partial H(Y_t)}{\partial t} &= -\mathbb{E} \left[\lambda \int_{-\infty}^{\infty} (\log P_t(Y_t + x) - \log P_t(Y_t)) \nu(dx) \right] \\ &= -\mathbb{E} \left[\lambda \int_0^{\infty} ((t-1)(\log(Y_t + x) - \log Y_t) - \lambda x) \frac{e^{-\lambda x}}{x} dx \right] \\ &= \lambda + \lambda(1-t) \mathbb{E} \left[\int_0^{\infty} (\log(Y_t + x) - \log Y_t) \frac{e^{-\lambda x}}{x} dx \right].\end{aligned}$$

by using Lemma 6.4.1. □

6.5 Log-Sobolev type inequalities for the relative entropy

In this section we extend the logarithmic Sobolev inequality in [41] (see 1.6.7) to the relative entropy of two probability measures on the Poisson space.

First we want to state the definition of the relative entropy for probability measures:

Definition 6.5.1. Let $\mathcal{P}(\Omega)$ denote the set of all probability measures defined on $(\Omega, \mathcal{F})$. For $\gamma \in \mathcal{P}(\Omega)$, the relative entropy function $D(\cdot||\gamma)$ is a mapping from $\mathcal{P}(\Omega)$ into $\mathbb{R} \cup \{\infty\}$ given by

$$D(\gamma'||\gamma) := \mathbb{E}^{\gamma'} \left[\log \frac{d\gamma'}{d\gamma} \right],$$

whenever $\gamma' \in \mathcal{P}(\Omega)$ is absolutely continuous with respect to γ such that the above integral is finite, where $\mathbb{E}^{\gamma'}$ denotes the expectation with respect to γ' . In all other cases, $D(\gamma'||\gamma) := \infty$.

Let γ' and γ be two probability measures on $(\Omega, \mathcal{F})$ whose densities are represented by

$$d\gamma' = \mathcal{E}_1(\phi)d\mathbb{P} \text{ and } d\gamma = \mathcal{E}_1(\tilde{\phi})d\mathbb{P} \quad (6.5.1)$$

where $\mathcal{E}_1(\cdot)$ is defined in (1.6.2) for bounded predictable processes ϕ and $\tilde{\phi}$.

Then we have the following expression for the relative entropy.

Lemma 6.5.2. Let γ', γ be two probability measures on $(\Omega, \mathcal{F})$, absolutely continuous with respect to $\mathbb{P}$, and whose densities are represented as in (6.5.1). Then we have

$$D(\gamma'||\gamma) = \mathbb{E}^{\gamma'} \left[\int_0^1 \int_{\mathbb{U}} \Psi(1 + \tilde{\phi}, \phi - \tilde{\phi}) \pi(du, ds) \right],$$

where

$$\Psi(x, y) = (x + y) \log(x + y) - (x + y) \log x - y, \quad x > 0, \quad x + y > 0.$$

Proof. Noting that

$$\mathbb{E}^{\gamma'} \left[\int_0^1 \int_{\mathbb{U}} \psi_{u,s} \mu(du, ds) \right] = \mathbb{E}^{\gamma'} \left[\int_0^1 \int_{\mathbb{U}} \psi_{u,s} (1 + \phi(u, s)) \pi(du, ds) \right],$$

for ψ a predictable process, we will have

$$\begin{aligned} D(\gamma' || \gamma) &= \mathbb{E}^{\gamma'} \left[\log \left(\frac{d\gamma'}{d\gamma} \right) \right] \\ &= \mathbb{E}^{\gamma'} \left[\log \left(\frac{d\gamma'}{d\mathbb{P}} \right) \right] - \mathbb{E}^{\gamma'} \left[\log \left(\frac{d\gamma}{d\mathbb{P}} \right) \right] \\ &= \mathbb{E}^{\gamma'} [\log(\mathcal{E}_1(\phi))] - \mathbb{E}^{\gamma'} [\log(\mathcal{E}_1(\tilde{\phi}))] \\ &= \mathbb{E}^{\gamma'} \left[\int_0^1 \int_{\mathbb{U}} \log(1 + \phi(u, s)) \tilde{\mu}(du, ds) + \int_0^1 \int_{\mathbb{U}} (\log(1 + \phi(u, s)) - \phi(u, s)) \pi(du, ds) \right] \\ &\quad - \mathbb{E}^{\gamma'} \left[\int_0^1 \int_{\mathbb{U}} \log(1 + \tilde{\phi}(u, s)) \tilde{\mu}(du, ds) + \int_0^1 \int_{\mathbb{U}} (\log(1 + \tilde{\phi}(u, s)) - \tilde{\phi}(u, s)) \pi(du, ds) \right] \\ &= \mathbb{E}^{\gamma'} \left[\int_0^1 \int_{\mathbb{U}} ((1 + \phi(u, s)) \log(1 + \phi(u, s)) - \phi(u, s)) \pi(du, ds) \right] \\ &\quad - \mathbb{E}^{\gamma'} \left[\int_0^1 \int_{\mathbb{U}} ((1 + \phi(u, s)) \log(1 + \tilde{\phi}(u, s)) - \tilde{\phi}(u, s)) \pi(du, ds) \right] \\ &= \mathbb{E}^{\gamma'} \left[\int_0^1 \int_{\mathbb{U}} \left((1 + \phi)(\log(1 + \phi) - \log(1 + \tilde{\phi})) + \tilde{\phi} - \phi \right) \pi(du, ds) \right] \\ &= \mathbb{E}^{\gamma'} \left[\int_0^1 \int_{\mathbb{U}} \Psi(1 + \tilde{\phi}, \phi - \tilde{\phi}) \pi(du, ds) \right] \end{aligned}$$

where

$$\Psi(x, y) = (x + y) \log(x + y) - (x + y) \log x - y, \quad x > 0, \quad x + y > 0.$$

□

Note that we have

$$\begin{aligned}
 \Psi(x, y) &= (x + y) \log(x + y) - (x + y) \log x - y \\
 &= (x + y)(\log(x + y) - \log x) - y \\
 &= (x + y) \log \left(1 + \frac{y}{x}\right) - y \\
 &\leq (x + y) \frac{y}{x} - y \\
 &= \frac{y^2}{x}, \quad x > 0, \quad x + y > 0. \tag{6.5.2}
 \end{aligned}$$

By using Lemma 6.5.2. we obtain the following theorem which extends the logarithmic Sobolev inequality in Theorem 1.1 of [41] to the relative entropy of two probability measures on the Poisson space.

Theorem 6.5.3. *Let γ', γ be two probability measures on $(\Omega, \mathcal{F})$, absolutely continuous with respect to $\mathbb{P}$, and whose densities are represented as in (6.5.1). We have the logarithmic Sobolev inequality*

$$D(\gamma' || \gamma) \leq \mathbb{E}^{\gamma'} \left[\int_0^1 \int_{\mathbb{U}} \frac{|\phi - \tilde{\phi}|^2}{1 + \tilde{\phi}} \pi(du, ds) \right].$$

Proof. By using the inequality in (6.5.2) we will have

$$D(\gamma' || \gamma) = \mathbb{E}^{\gamma'} \left[\int_0^1 \int_{\mathbb{U}} \Psi(1 + \tilde{\phi}, \phi - \tilde{\phi}) \pi(du, ds) \right] \leq \mathbb{E}^{\gamma'} \left[\int_0^1 \int_{\mathbb{U}} \frac{|\phi - \tilde{\phi}|^2}{1 + \tilde{\phi}} \pi(du, ds) \right].$$

□

By the Csiszár-Kullback-Pinsker inequality we find

$$\|\gamma' - \gamma\|_{TV}^2 \leq \frac{1}{2} D(\gamma' || \gamma) \leq \frac{1}{2} \mathbb{E}^{\gamma'} \left[\int_0^1 \int_{\mathbb{U}} \frac{|\phi - \tilde{\phi}|^2}{1 + \tilde{\phi}} \pi(du, ds) \right],$$

where

$$\|\gamma' - \gamma\|_{TV} = \frac{1}{2} \mathbb{E}^\gamma \left[\left| 1 - \frac{d\gamma'}{d\gamma} \right| \right] = \frac{1}{2} \mathbb{E} \left[\left| \frac{d\gamma}{d\mathbb{P}} - \frac{d\gamma'}{d\mathbb{P}} \right| \right] = \frac{1}{2} \mathbb{E} \left[\left| \frac{d\gamma}{d\mathbb{P}} - \frac{d\gamma'}{d\mathbb{P}} \right| \right].$$

Given F a non-negative random variable such that $E[F] = 1$, and γ' be the probability measure which has density F with respect to $\mathbb{P}$. Then the entropy $\text{Ent}[F]$ of F satisfies

$$\text{Ent}[F] = E[F \log F] = D(\gamma' || \mathbb{P}).$$

Taking $\gamma = \mathbb{P}$ and $\tilde{\phi} = 0$ in Theorem 6.5.3 then we will have the following logarithmic Sobolev inequality

$$\text{Ent}[F] = D(\gamma' || \mathbb{P}) \leq \frac{1}{2} \mathbb{E}^{\gamma'} [|\phi|^2] = \frac{1}{2} \mathbb{E}^{\gamma'} [|\phi|_{L^2(\pi)}^2] \quad (6.5.3)$$

for γ' a probability measure satisfying (6.5.1). This will help us to recover the logarithmic Sobolev inequality in Theorem 1.1 of [41] on the Poisson space. In the following theorem we recover the result in Theorem 1.1 of [41] with a different proof.

Theorem 6.5.4. *For any $F > 0$ a.s. we have*

$$\text{Ent}[F] \leq \frac{1}{2} \mathbb{E} \left[\frac{1}{F} \|DF\|_{L^2(\pi)}^2 \right].$$

Proof. By Proposition 4.2 in [43] the process ϕ in the representation

$$F = \frac{d\gamma'}{d\mathbb{P}} = \mathcal{E}_1(\phi)$$

is given by

$$\phi(u, s) = \frac{\mathbb{E}[D_{u,s}F \mid \mathcal{F}_t]}{\mathbb{E}[F \mid \mathcal{F}_t]}, \quad (u, s) \in \mathbb{U} \times [0, 1], \quad (6.5.4)$$

hence from (6.5.3) we have

$$\begin{aligned} D(\gamma' \parallel \gamma) &\leq \frac{1}{2} \mathbb{E}^{\gamma'} [\|\phi\|_{L^2(\pi)}^2] = \frac{1}{2} \mathbb{E}^{\gamma'} \left[\int_0^1 \int_{\mathbb{U}} |\phi(u, s)|^2 \pi(du, ds) \right] \\ &= \frac{1}{2} \mathbb{E}^{\gamma'} \left[\int_0^1 \int_{\mathbb{U}} \left| \frac{\mathbb{E}[D_{u,s}F \mid \mathcal{F}_s]}{\mathbb{E}[F \mid \mathcal{F}_s]} \right|^2 \pi(du, ds) \right] \\ &\leq \frac{1}{2} \mathbb{E}^{\gamma'} \left[\int_0^1 \int_{\mathbb{U}} \mathbb{E} \left[\left| \frac{D_{u,s}F}{F} \right|^2 \mid \mathcal{F}_s \right] \pi(du, ds) \right] \\ &= \frac{1}{2} \mathbb{E}^{\gamma'} \left[\int_0^1 \int_{\mathbb{U}} \left| \frac{D_{u,s}F}{F} \right|^2 \pi(du, ds) \right] \\ &= \frac{1}{2} \mathbb{E} \left[\frac{1}{F} \int_0^1 \int_{\mathbb{U}} |D_{u,s}F|^2 \pi(du, ds) \right] \end{aligned}$$

where we have applied the Jensen's inequality to the convex function $(x, y) \mapsto x^2/y^2$ on $\mathbb{R} \times (0, \infty)$ and the tower property of conditional expectation. $\square$

6.6 Calculations for predictable decompositions

Recall that the standard Poisson process is a stochastic process $(N_t)_{t \in \mathbb{R}_+}$ with jumps of size +1 only, and whose paths are constant in between two jumps, i.e. at time t , the value N_t of the process is given by

$$N_t = \sum_{k=1}^{\infty} 1_{[T_k, \infty)}(t), \quad t \in \mathbb{R}_+, \quad (6.6.1)$$

where

$$1_{[T_k, \infty)}(t) = \begin{cases} 1 & \text{if } t \geq T_k, \\ 0 & \text{if } 0 \leq t < T_k \end{cases} \quad (6.6.2)$$

$k \geq 1$, and $(T_k)_{k \geq 1}$ is the increasing family of jump times of $(N_t)_{t \in \mathbb{R}_+}$ such that

$$\lim_{k \rightarrow \infty} T_k = +\infty.$$

Furthermore we have the following equivalence relation

$$\{N_t \leq k\} \Leftrightarrow \{T_{k+1} > t\},$$

for all $k \geq 1$. In above T_{k+1} follows Gamma distribution with the density function

$$t \mapsto \lambda^{k+1} e^{-\lambda t} \frac{t^k}{k!}$$

for all $k \geq 0$ (see Proposition 15.2 of Privault [32]).

In this section we compute the predictable process

$$\phi(u, s) = \frac{\mathbb{E}[D_{(u,s)} F | \mathcal{F}_s]}{\mathbb{E}[F | \mathcal{F}_s]}, \quad (u, s) \in \mathbb{U} \times [0, 1] \quad (6.6.3)$$

which we used in Section 6.5 by using the Clark-Ocone formula.

In order to compute (6.6.3) we follow some notations and techniques from the paper of Mensi and Privault [20].

Let

$$p_n^t = \mathbb{P}(N_t = n) = 1_{\{n \geq 0\}} \frac{(t \vee 0)^n}{n!} e^{-t \vee 0}, \quad n \in \mathbf{Z}, t \in \mathbf{R}$$

with the convention $0^0 = 1$.

We apply the Itô formula

$$\begin{aligned} f(N_v, v) &= f(N_u, u) + \int_u^v (f(1 + N_{s-}, s) - f(N_{s-}, s)) d\tilde{N}_s \\ &\quad + \int_u^v (f(1 + N_s, s) - f(N_s, s)) ds + \int_u^v \frac{\partial f}{\partial s}(N_s, s) ds \end{aligned}$$

to the martingale in t

$$\begin{aligned} f(N_t(w), t) &= \mathbb{E}[1_{\{N_T - N_u = n\}} | \mathcal{F}_t] \\ &= \mathbb{E}[1_{\{N_T^t(w, \cdot) - N_u^t(w, \cdot) = n\}}] \\ &= \mathbb{E}[1_{\{N_{T-t}(\cdot) + N_t(w) - N_u(w) = n\}}] \\ &= p_{n-(N_t - N_u)(w)}^{T-t}, \quad u < t < T, \end{aligned}$$

which gives the predictable representation

$$p_{n-(N_v - N_u)}^{T-v} = p_n^{T-u} + \int_u^v (p_{n-1-N_t+N_u}^{T-t} - p_{n-N_t+N_u}^{T-t}) d\tilde{N}_t, \quad 0 \leq u \leq v \leq T$$

and in particular for $T = v$,

$$1_{\{N_v - N_u = n\}} = p_n^{v-u} + \int_u^v (p_{n-1-N_t+N_u}^{v-t} - p_{n-N_t+N_u}^{v-t}) d\tilde{N}_t, \quad n \in \mathbf{Z}.$$

For $F = 1_{\{N_v - N_u = n\}}$ and $u \leq t \leq v$ the above integrand is

$$\begin{aligned} p_{n-1-N_t+N_u}^{v-t} - p_{n-N_t+N_u}^{v-t} &= \mathbb{E}[1_{\{N_v - N_u + 1 = n\}} - 1_{\{N_v - N_u = n\}} | \mathcal{F}_t] \\ &= \mathbb{E}[D_t 1_{\{N_v - N_u = n\}} | \mathcal{F}_t] \end{aligned} \tag{6.6.4}$$

where D is the finite difference operator defined as

$$D_t f(N_v - N_u) = 1_{[u,v]}(t)(f(N_v - N_u + 1) - f(N_v - N_u)).$$

Also from (iii) of [20] we have

$$\begin{aligned} \mathbb{P}(T_n \geq x | \mathcal{F}_t)(w) &= \mathbb{E}[1_{\{T_n \geq x\}} | \mathcal{F}_t](w) \\ &= \mathbb{E}[1_{\{T_n^t(w, \cdot) \geq x\}}] \\ &= 1_{\{x \leq T_n(w) \leq t\}} + 1_{\{T_n(w) > t\}} \mathbb{E}[1_{\{t + T_{n-N_t(w)} \geq x\}}]. \end{aligned} \quad (6.6.5)$$

Proposition 6.6.1. *Let μ be a probability measure on $(\Omega, \mathcal{F}^{\mathbb{P}})$ defined by*

$$d\mu = \mathcal{E}_1(\phi) d\mathbb{P} = \frac{1_{\{N_v \leq k\}}}{\mathbb{P}(\{N_v \leq k\})} d\mathbb{P}$$

then we have

$$\phi(t) = - \frac{1}{\sum_{n=0}^{k-N_t} \binom{k-N_t}{n} (k-N_t-n)!(v-t)^{-(k-N_t-n)}} 1_{[0,v]}(t). \quad (6.6.6)$$

Proof. For $1_{\{N_v \leq k\}} = 1_{\{T_{k+1} > v\}}$ from (6.6.4) we have

$$\begin{aligned} \mathbb{E}[D_t 1_{\{N_v \leq k\}} | \mathcal{F}_t] &= \mathbb{P}(N_{v-t} \leq k-1-N_t) - \mathbb{P}(N_{v-t} \leq k-N_t) \\ &= 1_{[0,v]}(t) \sum_{n=0}^{k-1-N_t} e^{-(v-t)} \frac{(v-t)^n}{n!} - 1_{[0,v]}(t) \sum_{n=0}^{k-N_t} e^{-(v-t)} \frac{(v-t)^n}{n!} \\ &= -1_{[0,v]}(t) e^{-(v-t)} \frac{(v-t)^{k-N_t}}{(k-N_t)!} \end{aligned} \quad (6.6.7)$$

and from (6.6.5) we have

$$\begin{aligned}
 \mathbb{E}[1_{\{N_v \leq k\}} | \mathcal{F}_t] &= \mathbb{E}[1_{\{T_{k+1} > v\}} | \mathcal{F}_t] \\
 &= 1_{\{v \leq T_{k+1} \leq t\}} + 1_{\{T_{k+1} > t\}} \mathbb{E}[1_{\{t + T_{k+1} - N_t \geq v\}}] \\
 &= 1_{\{v \leq T_{k+1} \leq t\}} + 1_{\{T_{k+1} > t\}} \mathbb{E}[1_{\{N_{v-t} \leq k - N_t\}}] \\
 &= 1_{\{v \leq T_{k+1} \leq t\}} + 1_{\{T_{k+1} > t\}} \sum_{n=0}^{k-N_t} e^{-(v-t)} \frac{(v-t)^n}{n!}. \quad (6.6.8)
 \end{aligned}$$

Since in (6.6.7) we should have $\{N_t \leq k\}$ which is equivalent to $\{T_{k+1} > t\}$ then we will only take the second term in (6.6.8).

Then from (6.6.3) we will have,

$$\begin{aligned}
 \phi(t) &= \frac{\mathbb{E}[D_t 1_{\{N_v \leq k\}} | \mathcal{F}_t]}{\mathbb{E}[1_{\{N_v \leq k\}} | \mathcal{F}_t]} \\
 &= - \frac{1}{\sum_{n=0}^{k-N_t} \binom{k-N_t}{n} (k-N_t-n)! (v-t)^{-(k-N_t-n)}} 1_{[0,v]}(t).
 \end{aligned}$$

□

Proposition 6.6.2. *Let μ be a probability measure on $(\Omega, \mathcal{F}^{\mathbb{P}})$ defined by*

$$d\mu = \mathcal{E}_1(\phi) d\mathbb{P} = \frac{1_{\{N_v = k\}}}{\mathbb{P}(\{N_v = k\})} d\mathbb{P}$$

then we have

$$\phi(t) = - \frac{((v-t) - (k - N_t))}{(v-t)} 1_{[0,v]}(t). \quad (6.6.9)$$

Proof. From (6.6.4) we have

$$\mathbb{E}[D_t 1_{\{N_v = k\}} | \mathcal{F}_t] = \mathbb{P}(N_{v-t} = k - 1 - N_t) - \mathbb{P}(N_{v-t} = k - N_t)$$

$$\begin{aligned}
 &= 1_{[0,v]}(t)e^{-(v-t)}\frac{(v-t)^{k-1-N_t}}{(k-1-N_t)!} - 1_{[0,v]}(t)e^{-(v-t)}\frac{(v-t)^{k-N_t}}{(k-N_t)!} \\
 &= -1_{[0,v]}(t)e^{-(v-t)}(v-t)^{k-N_t}\frac{((v-t)-(k-N_t))}{(v-t)(k-N_t)!}.
 \end{aligned}$$

By using the fact $1_{\{N_v=k\}} = 1_{\{N_v \leq k\}} - 1_{\{N_v \leq k-1\}}$ from (6.6.8) we will have

$$\begin{aligned}
 \mathbb{E}[1_{\{N_v=k\}}|\mathcal{F}_t] &= \mathbb{E}[1_{\{N_v \leq k\}}|\mathcal{F}_t] - \mathbb{E}[1_{\{N_v \leq k-1\}}|\mathcal{F}_t] \\
 &= \sum_{n=0}^{k-N_t} e^{-(v-t)}\frac{(v-t)^n}{n!} - \sum_{n=0}^{k-N_t-1} e^{-(v-t)}\frac{(v-t)^n}{n!} \\
 &= e^{-(v-t)}\frac{(v-t)^{k-N_t}}{(k-N_t)!}.
 \end{aligned}$$

Then we have

$$\begin{aligned}
 \phi(t) &= \frac{\mathbb{E}[D_t 1_{\{N_v=k\}}|\mathcal{F}_t]}{\mathbb{E}[1_{\{N_v=k\}}|\mathcal{F}_t]} \\
 &= -\frac{((v-t)-(k-N_t))}{(v-t)} 1_{[0,v]}(t).
 \end{aligned}$$

□

Remark 6.6.3 (No Jump Case). *Let μ be the probability measure on $(\Omega, \mathcal{F}^{\mathbb{P}})$ defined by*

$$d\mu = \mathcal{E}_1(\phi)d\mathbb{P} = \frac{1_{\{N_v=0\}}}{\mathbb{P}(\{N_v=0\})}d\mathbb{P}$$

then we have

$$\phi(t) = -1_{[0,v]}(t), \quad 0 \leq t \leq v. \quad (6.6.10)$$

Future Research

In this chapter we present several open problems which may be a topic of interest for future research.

Supermodular ordering of Poisson random vectors

Supermodular ordering of Gaussian random vectors has been completely characterized in [25] Theorem 4.2 and cf. also Theorem 3.13.5 of [26] (see Theorem 1.1.7), by showing that

- (i) X_i and Y_i have same distribution for all $i = 1, \dots, d$, and
- (ii) for all $1 \leq i < j \leq d$ we have

$$\text{Cov}(X_i, X_j) \leq \text{Cov}(Y_i, Y_j),$$

are necessary and sufficient conditions for supermodular ordering of Gaussian random vectors X and Y .

Then, naturally, we ask what are necessary and sufficient conditions for supermodular ordering of Poisson random vectors? Since conditions (i) and (ii) above are necessary conditions for supermodular ordering of Poisson random vectors then are they also sufficient? In [18] (see Theorem 2.4.2), we showed that (i) and (ii) above are also sufficient conditions for a particular dependence structure of Poisson random vectors depending on $d(d+1)/2$ parameters based on a natural decomposition of their

Lévy measure on the vertices of the d -dimensional unit hypercube. Later, in [17] (see Theorem 3.4.2) we extended our previous result to a much larger family of dependence structures; by using binary tree and using Möbius inversion techniques to show that (i) and (ii) above are also sufficient conditions for supermodular ordering of Poisson random vectors. There is, still, a gap in this area of research: we need to find all possible dependence structures, other than tree structure, for Poisson random vectors such that (i) and (ii) to become necessary and sufficient for supermodular ordering of Poisson random vectors.

Co-monotonic ordering of random vectors

The definition of co-monotonic functions is given in the Definition 5.4 of Molchanov [22] in the following way:

Definition 6.6.4 (Co-monotonic functions). *Real-valued functions f and g are called co-monotonic if*

$$(f(x) - f(x'))(g(x) - g(x')) \geq 0 \text{ for all } x, x' \in \mathbb{R}^d. \quad (6.6.11)$$

A possible definition for co-monotonic ordering of random vectors can be stated in the following way

Definition 6.6.5. *Let X and Y be d -dimensional random vectors. Then X and Y are comparable with respect to co-monotonic ordering if*

$$\text{Cov}(f(X), g(X)) \leq \text{Cov}(f(Y), g(Y)) \quad (6.6.12)$$

for all co-monotonic functions f and g .

Now, naturally, one can ask what are necessary and sufficient conditions for ‘co-monotonic’ ordering of Gaussian and Poisson random vectors? By using the Definition 6.6.5, it is not easy to derive necessary and sufficient conditions for ‘co-monotonic’ ordering of Gaussian and Poisson random vectors; then what are possible definitions for ‘co-monotonic’ ordering of random vectors so that we can characterize necessary and sufficient conditions for Gaussian and Poisson random vectors?

Propagation of supermodularity

Consider the following multidimensional drift-diffusion-jump model

$$dX(t) = \alpha(X(t), t)dt + \beta(X(t), t)dW(t) + \int_0^1 \phi(X(t), t, z)\tilde{\nu}(dt, dz) \quad (6.6.13)$$

where $X(t)$ is a d -dimensional stochastic process, $\alpha = (\alpha_i)_{i=1}^d$ is a d -dimensional vector, $\beta = (\beta_{ij})_{i,j=1}^d$ is a $d \times d$ -matrix, $W(t)$ is a d -dimensional Brownian motion, $\phi = (\phi_1, \dots, \phi_d)$ is a d -dimensional vector, and $\tilde{\nu}$ is the compensated jump measure defined by $\tilde{\nu}(dt, dz) = (\nu - q)(dt, dz)$ such that ν is a Poisson random measure on $[0, T] \times [0, 1]$ with intensity measure $q(dt, dz) = \lambda(t)dtdz$ where λ is a deterministic function.

Let $f : \mathbb{R}^d \rightarrow \mathbb{R}$ be a supermodular function then under which conditions of α , β , and ϕ

$$\mathcal{F}(t, x) = \mathbb{E}[f(X(T)) | X(t) = x] \quad (6.6.14)$$

is a supermodular function? We want to find out which models are supermodularity preserving so that we can establish sufficient conditions to derive bounds for option prices.

In Theorem 2.6 of [5] and in Theorem 3.2 and Theorem 3.6 of [4] propagation of convexity and directionally convexity for some multidimensional examples have been proved. In Proposition 4.1 of Ekström and Tysk [11] and in Theorem 4.2 of Ekström and Tysk [10], under some regularity conditions, it is shown that a model is convexity preserving if and only if the operator (which is written by using drift, diffusion, and jump part of the model) is locally convexity preserving. To our best knowledge, there is no results for propagation of supermodularity.

Entropy of multidimensional Poisson distribution

The differential entropy of multidimensional Gaussian distribution can be calculated easily by using its density function however it is not easy to calculate the entropy of multidimensional Poisson distribution; there are only some approximations in the literature, see for example Section 5 of Guerrero-Cusumano [13]. The entropy of multidimensional Poisson distribution is still an open question.

Conclusion

Necessary and sufficient conditions for supermodular and convex ordering of Gaussian random vectors have been characterized in [25] and [26]. Propagation of convexity and sufficient conditions for comparison of European options with respect to martingale measures is obtained in [4] and [5]. Entropy of a multivariate Gaussian distribution and bounds for information theory measures are derived in [9]. In this thesis, first of all, we derive necessary and sufficient conditions for supermodular ordering of Poisson random vectors of a certain structure and then we extend our results by using binary trees and Möbius inversion techniques. Moreover, we obtain the same necessary and sufficient conditions for supermodular ordering of Binomial random vectors and in the limit case we recover supermodular ordering of Poisson random vectors and existing results in the literature for Gaussian random vectors. Secondly, we prove propagation of supermodularity for multidimensional Ornstein-Uhlenbeck process and multidimensional jump-diffusion model. Furthermore, we derive sufficient conditions for bounding option prices for multidimensional Ornstein-Uhlenbeck process by using increasing supermodular ordering. Lastly, we calculate information theory measures on the Wiener and on the Poisson spaces by using Itô calculus and we derive bounds for information theory measures.

Bibliography

- [1] APPLEBAUM, D. *Lévy Processes and Stochastic Calculus*. Cambridge Studies in Advanced Mathematics. Cambridge University Press, 2004.
- [2] ARNAUDON, M., BRETON, J.-C., AND PRIVAULT, N. Convex Ordering for Random Vectors using Predictable Representation. *Potential Analysis* 29(4) (2008), 327–349.
- [3] BÄUERLE, N., MÜLLER, A., AND BLATTER, A. Dependence properties and comparison results for Lévy processes. *Math. Meth. Operat. Res.* 67 (2008), 161–186.
- [4] BERGENTHUM, J., AND RÜSCHENDORF, L. Comparison of Option Prices in Semimartingale Models. *Finance and Stochastics* 10(2) (2006), 222–249.
- [5] BERGENTHUM, J., AND RÜSCHENDORF, L. Comparison of semimartingales and Lévy processes. *Annals of Probability* 35 (2007), 228–254.

REFERENCES

- [6] BESAL, M., KOHATSU-HIGA, A., AND TINDEL, S. Gaussian-type lower bounds for the density of solutions of sdes driven by fractional brownian motions. *Ann. Probab.* *44*, 1 (01 2016), 399–443.
- [7] BOBKOV, S., AND GÖTZE, F. Exponential Integrability and Transportation Cost Related to Logarithmic Sobolev Inequalities . *Journal of Functional Analysis* *163*, 1 (1999), 1–28.
- [8] CICALESE, F., AND VACCARO, U. Supermodularity and subadditivity properties of the entropy on the majorization lattice. *IEEE Trans. Information Theory* *48*, 4 (2002), 933–938.
- [9] COVER, T. M., AND THOMAS, J. A. *Elements of Information Theory*. Wiley-Interscience, New York, NY, USA, 1991.
- [10] EKSTRÖM, E., AND TYSK, J. Convexity preserving jump-diffusion models for option pricing. *Journal of Mathematical Analysis and Applications* *330*(1) (2007), 715–728.
- [11] EKSTRÖM, E., AND TYSK, J. Properties of option prices in models with jumps. *Mathematical Finance* *17*(3) (2007), 381–397.
- [12] ES-SEBAIY, K., AND TUDOR, C. A. Multidimensional bifractional Brownian motion: Itô and Tanaka formulas. *Stochastics and Dynamics* *07*, 03 (2007), 365–388.
- [13] GUERRERO-CUSUMANO, J.-L. The entropy of the multivariate Poisson: An approximation . *Information Sciences* *86*, 13 (1995), 1–17.

REFERENCES

- [14] HOUDRÉ, C., PÉREZ-ABREU, V., AND SURGAILIS, D. Interpolation, correlation identities, and inequalities for infinitely divisible variables. *Journal of Fourier Analysis and Applications* 4, 6 (1998), 651–668.
- [15] ISHIKAWA, Y. *Stochastic calculus of variations*, second ed., vol. 54 of *De Gruyter Studies in Mathematics*. De Gruyter, Berlin, 2016. For jump processes.
- [16] KAWAMURA, K. The structure of multivariate Poisson distribution. *Kodai Math. J.* 2, 3 (1979), 337–345.
- [17] KIZILDEMIR, B., AND PRIVAULT, N. Supermodular ordering of binomial, Poisson and Gaussian random vectors by tree-based correlations. <http://arxiv.org/abs/1512.08883>.
- [18] KIZILDEMIR, B., AND PRIVAULT, N. Supermodular ordering of Poisson arrays . *Statistics and Probability Letters* 98 (2015), 136–143.
- [19] MAO, X. *Stochastic Differential Equations and Applications*, 2nd ed. Horwood Pub., 2008.
- [20] MENSI, M., AND PRIVAULT, N. Conditional calculus on Poisson space and enlargement of filtration. *Stochastic Analysis and Applications* 21, Issue 1 (2003), 183–204.
- [21] MEYER, M., AND STRULOVICI, B. The Supermodular Stochastic Ordering. Tech. Rep. 655, University of Oxford, Department of Economics, 2013.

REFERENCES

- [22] MOLCHANOV, I. *Theory of Random Sets*. Probability and Its Applications. Springer, 2005.
- [23] MÜLLER, A. Stop-loss order for portfolios of dependent risks . *Insurance: Mathematics and Economics* 21, 3 (1997), 219–223.
- [24] MÜLLER, A. Stochastic Ordering of Multivariate Normal Distributions. *Annals of the Institute of Statistical Mathematics* 53(3) (2001), 567–575.
- [25] MÜLLER, A., AND SCARSINI, M. Some Remarks on the Supermodular Order. *Journal of Multivariate Analysis* 73, 1 (2000), 107–119.
- [26] MÜLLER, A., AND STOYAN, D. *Comparison Methods for Stochastic Models and Risks*. Wiley Series in Probability and Statistics. Wiley, 2002.
- [27] NARASIMHAN, M., AND BILMES, J. A submodular-supermodular procedure with applications to discriminative structure learning. In *Proceedings of the Twenty-First Conference Annual Conference on Uncertainty in Artificial Intelligence (UAI-05)* (Arlington, Virginia, 2005), AUAI Press, pp. 404–412.
- [28] NGUYEN, T. D., PRIVAULT, N., AND TORRISI, G. L. Gaussian estimates for the solutions of some one-dimensional stochastic equations. *Potential Analysis* 43, 2 (2015), 289–311.
- [29] NOURDIN, I., AND VIENS, F. Density Formula and Concentration Inequalities with Malliavin Calculus. *Electron. J. Probab.* 14 (2009), no. 78, 2287–2309.

REFERENCES

- [30] ORMOS, M., AND ZIBRICZKY, D. Entropy-Based Financial Asset Pricing. *PLoS ONE* 9, (12) (December 2014), e115742.
- [31] PECCATI, G., AND TAQQU, S. M. *Wiener Chaos: Moments, Cumulants and Diagrams: A survey with Computer Implementation*. Bocconi & Springer Series. Springer Milan, 2011.
- [32] PRIVAULT, N. *Stochastic Finance: An Introduction with Market Examples*. Chapman & Hall/CRC Financial Mathematics Series. Taylor & Francis, 2013.
- [33] ROTA, G.-C. On the foundations of combinatorial theory. I. Theory of Möbius functions. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* 2 (1964), 340–368.
- [34] SATO, K. *Lévy Processes and Infinitely Divisible Distributions*. Cambridge University Press, Cambridge, 1999.
- [35] SHAKED, M., AND SHANTHIKUMAR, G. *Stochastic Orders*. Springer Series in Statistics. Springer New York, 2007.
- [36] SHANNON, C. E. A Mathematical Theory of Communication. *The Bell System Technical Journal* 27, 3 (1948), 379–423.
- [37] SIM, C. Generation of Poisson and Gamma random vectors with given marginals and covariance matrix. *Journal of Statistical Computation and Simulation* 47, 1-2 (1993), 1–10.
- [38] STUTZER, M. J. Simple Entropic Derivation of a Generalized Black-Scholes Option Pricing Model. *Entropy* 2, 2 (2000), 70–77.

REFERENCES

- [39] T. HU, C. X., AND RUAN, L. Dependence structures of multivariate Bernoulli random vectors. *J. Multivariate Anal.* *94(1)* (2005), 172–195.
- [40] TEICHER, H. On the multivariate Poisson distribution. *Scandinavian Actuarial Journal* *1954*, 1 (1954), 1–9.
- [41] WU, L. A new modified logarithmic Sobolev inequality for Poisson point processes and several applications. *Probab. Theory Related Fields* *118* (2000), 427–438.
- [42] YEUNG, R. W. *Information Theory and Network Coding*, 1 ed. Springer Publishing Company, Incorporated, 2008.
- [43] ZHANG, X. Clark-Ocone formula and variational representation for Poisson functionals. *Ann. Probab.* *37*, Number 2 (2009), 506–529.
- [44] ZHOU, R., CAI, R., AND TONG, G. Applications of Entropy in Finance: A Review. *Entropy* *15*, 11 (2013), 4909–4931.