นอร์มที่สอดคล้องและการวางนัยทั่วไปของผลคูณ * สำหรับคอปูลา

นาย พงศ์พล เรือนคง

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต
สาขาวิชาคณิตศาสตร์ ภาควิชาคณิตศาสตร์
คณะวิทยาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
ปีการศึกษา 2553

A COMPATIBLE NORM AND A GENERALIZATION

OF THE $*$-PRODUCT FOR COPULAS

Mr. Pongpol Ruankong

A Thesis Submitted in Partial Fulfillment of the Requirements

for the Degree of Master of Science Program in Mathematics

Department of Mathematics

Faculty of Science

Chulalongkorn University

Academic Year 2010

Thesis Title	A COMPATIBLE NORM AND A GENERALIZATION OF THE $*$-PRODUCT FOR COPULAS
By	Mr. Pongpol Ruankong
Field of Study	Mathematics
Thesis Advisor	Assistant Professor Songkiat Sumetkijakan, Ph.D.

Accepted by the Faculty of Science, Chulalongkorn University in Partial Fulfillment of the Requirements for the Master's Degree

.... Dean of the Faculty of Science

(Professor Supot Hannongbua, Dr.rer.nat.)

THESIS COMMITTEE

.... Chairman

(Associate Professor Wicharn Lewkeeratiyutkul, Ph.D.)

.... Thesis Advisor

(Assistant Professor Songkiat Sumetkijakan, Ph.D.)

.... Examiner

(Assistant Professor Nattakarn Chaidee, Ph.D.)

.... External Examiner

(Aram Tangboondouangjit, Ph.D.)

พงศ์พล เรือนคง : นอร์มที่สอดคล้องและการวางนัยทั่วไปของผลคูณ * สำหรับคอปูลา. (A COMPATIBLE NORM AND A GENERALIZATION OF THE *-PRODUCT FOR COPULAS) อ.ที่ปรึกษาวิทยานิพนธ์หลัก: ผศ.ดร.ทรงเกียรติ สุเมธกิจการ, 37 หน้า.

วิทยานิพนธ์นี้มีด้วยกัน 2 ส่วน ส่วนแรก เราสร้างนอร์มและแสดงว่ามันไม่แปรเปลี่ยนภายใต้การคูณ * ทางซ้ายและทางขวาด้วยคอปูลาที่มีอินเวอร์ส จากนั้นเราแสดงว่านอร์มนี้รักษาคุณสมบัติที่เกี่ยวข้องกับตัวแปรสุ่มบนเซ็ตของคอปูลา สำหรับการนำไปใช้ เราสร้างตัววัดความขึ้นต่อกันของตัวแปรสุ่มซึ่งไม่แปรเปลี่ยนภายใต้การแปลงตัวแปรสุ่มแบบโบเรล สำหรับส่วนที่สอง เราพิจารณาการวางนัยทั่วไปของผลคูณ * โดยเฉพาะการนิยามของมัน ซึ่งนิยามเป็นครั้งแรกโดย F. Durante, E.P. Klement และ J.J. Quesada-Molina นอกจากนั้นเรายังได้แสดงคุณสมบัติบางประการของมันอีกด้วย

ภาควิชา............คณิตศาสตร์.......... ลายมือชื่อนิสิต.....พงศ์พล เรือนคง.....
สาขาวิชา..........คณิตศาสตร์.......... ลายมือชื่อ อ.ที่ปรึกษาวิทยานิพนธ์หลัก.....ทรงเกียรติ สุเมธกิจการ.....
ปีการศึกษา.............2553...............

5272434123 : MAJOR MATHEMATICS
KEYWORDS : COPULA, $*$-PRODUCT, C-PRODUCT, SHUFFLE OF M, SHUFFLE OF COPULA, MEASURE OF DEPENDENCE

PONGPOL RUANKONG : A COMPATIBLE NORM AND A GENERALIZATION OF THE $*$-PRODUCT FOR COPULAS. ADVISOR: ASST.PROF. SONGKIAT SUMETKIJAKAN, Ph.D., 37 pp.

There are two parts in this thesis. In the first part, we define a norm and show that it is invariant under left and right $*$-products with invertible copulas. We then show that, restricted to the set of copulas, this norm preserves stochastic properties. As an application of the norm, we construct a measure of dependence which is invariant under Borel measurable bijective transformations. For the second part, we look into a generalization of the $*$-product, especially its definition, introduced by Durante, Klement and Quesada-Molina. We also derive some of its properties.

Department : ....Mathematics....	Student's Signature..............................
Field of Study : ....Mathematics....	Advisor's Signature..............................
Academic Year : ....2010...........	

ACKNOWLEDGEMENTS

I am greatly indebted to Assistant Professor Dr. Songkiat Sumetkijakan, my thesis advisor, for his suggestions and helpful advice in preparing and writing this thesis. I am also sincerely grateful to the thesis committee, for their suggestions and valuable comments. Moreover, I would like to thank all of my teachers who have taught me. I would like to express my deep gratitude to my beloved family for their love and encouragement.

Finally, I would like to thank all my friends for giving me invaluable experiences at Chulalongkorn university.

CONTENTS

CHAPTER I

INTRODUCTION AND PRELIMINARIES

1.1 Introduction

In 1992, for the purpose of studying Markov processes via copulas, Darsow, Nguyen and Olsen introduced a bilinear operation on the set of (2-dimensional) copulas known as the $*$-product. There are many more following researches on the $*$-product, e.g., invertible copulas with respect to the $*$-product, the joint continuity of the $*$-product with respect to various norms on the linear span of the set of copulas, some generalizations, etc. In this thesis, we focus on two aspects of the $*$-product: finding a "compatible" norm and studying a generalization.

In their paper, Darsow et al. showed that a copula of two random variables, which are conditionally independent given a third random variable, can be decomposed as a product of two copulas related to the three random variables. To be precise, if X and Y are conditionally independent given Z, then $C_{X,Y} = C_{X,Z} * C_{Z,Y}$, where $C_{X,Y}$ denotes a copula of the random vector (X, Y). We study a special case of the previous result where we transform the random variables X, Y. We obtain that if f and g are Borel measurable transformations of random variables X, Y, respectively, then $C_{f(X),g(Y)} = C_{f(X),X} * C_{X,Y} * C_{Y,g(Y)}$. In particular, if f and g are Borel measurable bijective transformations, Darsow et al. showed that copulas $C_{f(X),X}$ and $C_{Y,g(Y)}$ are invertible with respect to the $*$-product. To study $C_{f(X),g(Y)}$, we then study a more general form of the decomposition which we call *shuffling maps* on the linear span of the set of copulas: $A \mapsto U * A * V$ where

U, V are invertible. We discovered that, restricted to the set of copulas, shuffling maps preserve stochastic properties of copulas, i.e., they preserve independence, complete dependence and mutual complete dependence. In other words, the transformed random variables $f(X)$ and $g(Y)$ are independent, completely dependent or mutually completely dependent if and only if random variables X and Y are independent, completely dependent or mutually completely dependent, respectively. In the sense of this previous result, we can say that a suitable measure of dependence should then be invariant under bijective transformations. In order to obtain such a measure of dependence, we constructed a norm called the $*$-norm via $\|A\|_* = \sup\limits_{U,V \in \operatorname{Inv} \mathfrak{C}} \|U * A * V\|$ where $\operatorname{Inv} \mathfrak{C}$ denotes the set of invertible copulas and $\|\cdot\|$ denotes the Sobolev norm for copulas. We obtain that shuffling maps are isometies with respect to the norm. Then we construct the measure ω_* of two continuous random variables to be the normalized $*$-distance between the product copulas and the copula corresponding to the two random variables.

For the second half of the thesis, we study a generalization of the $*$-product known as $\mathbf{C}$-product. But to emphasize the link with the $*$-product, we will call it $*_{\mathbf{C}}$ product. This generalization arose from a research by Durante, Klement and Quesada-Molina on compatibility of copulas and characterizing Fréchet classes. For a family of copulas $\mathbf{C} = \{C_t\}_{t \in [0,1]}$, the $*_{\mathbf{C}}$ product of copulas A and B is given by

$$(A *_{\mathbf{C}} B)(x, y) = \int_0^1 C_t(\partial_2 A(x, t), \partial_1 B(t, y))\, dt.$$

However, it is questionable whether the product is well-defined because of the measurability of the integrand. In this part of our thesis, we restrict our attention to some reasonably large classes of families of copulas. Then, for each family $\mathbf{C}$ in those classes, we show that the $*_{\mathbf{C}}$ product is well-defined. Then, we derive some properties of the re-defined $*_{\mathbf{C}}$ product.

1.2 Preliminaries

In this section, we recall necessary definitions and properties involving our work. Here, we give a definition of bivariate copulas, or 2-copulas. We also give a definition of trivariate copulas, or 3-copulas, as we will encounter them later in a definition of classes of compatible copulas. Though, we are only interested in properties of 2-copulas. More details on copulas can be found in the classic book [7] by Nelsen.

Definition 1.1. A 2-copula, or simply copula, is a function $C\colon [0,1]^2 \to [0,1]$ satisfying the conditions:

1. $C(u,0) = C(0,v) = 0$ for all $u, v \in [0,1]$.
2. $C(u,1) = u$ and $C(1,v) = v$ for all $u, v \in [0,1]$.
3. C is 2-increasing, i.e., for all $[u_1,u_2] \times [v_1,v_2] \subseteq [0,1]^2$, we have

$$C(u_2,v_2) - C(u_2,v_1) - C(u_1,v_2) + C(u_1,v_1) \geq 0.$$

Definition 1.2. A 3-copula is a function $C\colon [0,1]^3 \to [0,1]$ satisfying the conditions:

1. $C(u,v,0) = C(u,0,w) = C(0,v,w) = 0$ for all $u, v, w \in [0,1]$.
2. $C(u,1,1) = u$, $C(1,v,1) = v$ and $C(1,1,w) = w$ for all $u, v, w \in [0,1]$.
3. C is 3-increasing, i.e., for all $[u_1,u_2] \times [v_1,v_2] \times [w_1,w_2] \in [0,1]^3$, we have

$$C(u_2,v_2,w_2) - C(u_1,v_2,w_2) - C(u_2,v_1,w_2) - C(u_2,v_2,w_1) +$$
$$C(u_2,v_1,w_1) + C(u_1,v_2,w_1) + C(u_1,v_1,w_2) - C(u_1,v_1,w_1) \geq 0.$$

According to Sklar's theorem (see, e.g., [7]), for any random vector (X, Y), there exists a copula C which links the joint distribution to its marginals as follows:

$$F_{XY}(u, v) = C(F_X(u), F_Y(v)).$$

If X and Y are continuous random variables, then the copula C is unique. We write $C_{X,Y}$ to represent a copula of the random vector (X, Y).

We denote the set of copulas by $\mathfrak{C}$. Every copula is Lipschitz continuous with Lipschitz constant 1, consequently, its partial derivatives exist almost everywhere and are bounded, wherever exist, between 0 and 1. Moreover, each copula induces a measure on the Borel subsets of $[0, 1]^2$ as follows.

Definition 1.3. Given a copula C, define a set function μ_C on the set of rectangles $[x_1, x_2] \times [y_1, y_2] \subseteq [0, 1]^2$ via

$$\mu_C([x_1, x_2] \times [y_1, y_2]) = C(x_2, y_2) - C(x_2, y_1) - C(x_1, y_2) + C(x_1, y_1) \geq 0.$$

Then, by standard measure-theoretic techniques, μ_C can be extended to a measure on the Borel σ-algebra on $[0, 1]^2$. Moreover, μ_C is doubly stochastic in the sense that $\mu_C(B \times [0, 1]) = \mu_C([0, 1] \times B) = \lambda(B)$ for every Borel set $B \subseteq [0, 1]$ where λ denotes Lebesgue measure. This measure is sometimes referred to as C-*measure*, C-*volume* or *mass of copula* C.

Definition 1.4. The support of a copula C, denoted by $\operatorname{supp} C$, is defined to be the complement of the union of all open subsets of $[0, 1]^2$ with zero C-volume.

The support of a copula C together with C-volume can be used to compute values of the copula at some, if not all, points $(x, y) \in [0, 1]^2$. We demonstrate such technique in the following example.

Example 1.5. Let C be a copula with its support shown in the figure below.

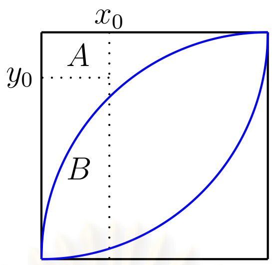

Figure 1.1: the support of C

For any point (x_0, y_0) in the upper left area, let A denote the rectangle whose vertices are $(0, y_0), (x_0, y_0), (x_0, 1)$ and $(0, 1)$ and let B denote the rectangle whose vertices are $(0, 0), (0, x_0), (x_0, y_0)$ and $(0, y_0)$. Then $\mu_C(A) = 0$ since it does not intersect the support of C. Moreover, $\mu_C(A \cup B) = C(x_0, 1) - C(x_0, 0) - C(0, 1) + C(0, 0) = x_0$. Then, $\mu_C(B) = \mu_C(A \cup B) - \mu_C(A) = x_0$. Hence,

$$C(x_0, y_0) = \mu_C(B) + C(0, y_0) + C(x_0, 0) - C(0, 0) = x_0.$$

Notice that the values of C at the points in the lower right area can be computed similarly.

Theoretically, the most important copulas are the Fréchet-Hoeffding upper and lower bounds and the product copula. The formulae are given, respectively, by

$$M(u, v) = \min(u, v),$$

$$W(u, v) = \max(u + v - 1, 0),$$

$$\Pi(u, v) = uv.$$

These copulas represent comonotonicity, countermonotonicity and independence, respectively, between the two random variables.

Example 1.6. It can be shown that $\operatorname{supp} M$ is the main diagonal from $(0,0)$ to $(1,1)$, $\operatorname{supp} W$ is the other diagonal and Π has full support, i.e., $\operatorname{supp} \Pi = [0,1]^2$.

In their study of Markov processes, Darsow, Nguyen and Olsen [1, p. 604] introduced a binary operation $*\colon \mathfrak{C} \times \mathfrak{C} \to \mathfrak{C}$ defined by

$$(A * B)(u,v) = \int_0^1 \partial_2 A(u,t) \partial_1 B(t,v)\ dt,$$

where ∂_i denotes the partial derivative with respect to the i-th variable. This operation is bilinear and is called the $*$-*product* on $\mathfrak{C}$. Remark that it can be naturally extended to a bilinear operation on $\operatorname{span} \mathfrak{C}$.

From straightforward computations, for any $C \in \mathfrak{C}$, we have the following identities: $M * C = C * M = C$ and $\Pi * C = C * \Pi = \Pi$. Therefore, copulas M and Π can be viewed as the identity and the zero of $(\mathfrak{C}, *)$, respectively. Moreover, denoted by C^T, the transpose of C, defined by $C^T(u,v) = C(v,u)$ is also a copula. In addition, a copula B is said to be invertible if there exists a copula C such that $B * C = C * B = M$. The set of invertible copulas plays an important role in this thesis and is denoted by $\operatorname{Inv} \mathfrak{C}$.

Remark 1.7. If they exist, left and right inverses of a copula $C \in \mathfrak{C}$ are unique and given by the transposed copula C^T (for a proof, see [1, Theorem 7.1]).

An important class of invertible copulas is the class of shuffles of M. This class attracts our interest because it is easy to compute. Moreover, Santiwipanont and Sumetkijakan [9] showed that the set of shuffles of M is dense in $\operatorname{Inv} \mathfrak{C}$ with respect to the Sobolev norm for copulas. A definition of a shuffle of M is given below. For more details on shuffles of M, see, e.g., [5, 9].

Definition 1.8. A copula C is a *shuffle of* M if and only if there exist a positive integer n, partitions $0 = s_0 < s_1 < \cdots < s_n = 1$ and $0 = t_0 < t_1 < \cdots < t_n = 1$ of $[0,1]$, and a permutation σ on the set $\{1, 2, \ldots, n\}$ such that each $(s_{i-1}, s_i) \times (t_{\sigma(i)-1}, t_{\sigma(i)})$ is a square of C-volume $s_i - s_{i-1}$ and its intersection with the support of C is one of the diagonals of the square. In this thesis, we call it a shuffle of M of n stripes.

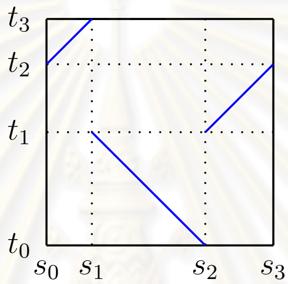

Figure 1.2: the support of a shuffle of M where $\sigma = (1\ 3\ 2)$

Example 1.9. The *straight shuffle of* M at $\alpha \in [0,1]$, denoted by S_α, is defined to be the shuffle of M supported on the straight line joining the points $(0,\alpha)$ and $(1-\alpha, 1)$ and the straight line joining the points $(1-\alpha, 0)$ and $(1, \alpha)$.

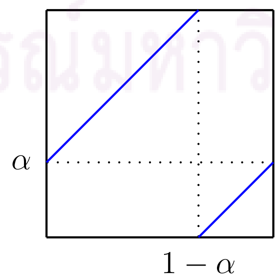

Figure 1.3: the support of the straight shuffle of M at $\alpha \in [0,1]$

Durante, Sarkoci and Sempi [5] generalized the idea of shuffles of M to shuffles of any copula. The definition of shuffles of copulas is measure theoric. Fortunately, Santiwipanont et al. [9] gave a useful characterization of shuffles of copulas: a shuffle of copula C is the $*$-product of the copula C with a shuffle of M on the left. They also introduce generalized shuffles of copulas: the $*$-product of a copula with an invertible copula on the left. This idea can be extended further, i.e., the $*$-product of a copula with two invertible copulas, one on the left and the other on the right, which we call *two-sided generalized shuffles of copulas.* For more details on shuffles of copulas, see [5, 9].

Example 1.10 ([9], p. 14)**.** Let S_α be the straight shuffle of M at $\alpha \in [0,1]$ as in Example 1.9. Then

$$(S_\alpha * C)(u,v) = \begin{cases} C(u+1-\alpha, v) - C(1-\alpha, v) & \text{if } \ 0 \le u \le \alpha \le 1 \\ v - C(1-\alpha, v) + C(u-\alpha, v) & \text{if } \ 0 \le \alpha \le u \le 1. \end{cases}$$

Observe that the mass of copula C is shuffled according to the shuffling of S. We generalize this observation in the following remark.

Remark 1.11. Let's remark that a shuffle of copula C, which can be written as $S * C$ where S is a shuffle of M, is indeed the shuffling of the mass of copula C according to the shuffling of S. In particular, the support of C is also shuffled accordingly. This fact can be shown by considering a set of generators of the set of shuffles of M, which is the set of all shuffles of M of three stripes where the first is fixed while the second and the third are swapped such that the swapped second stripe is straight, i.e., the support in that stripe has slope one. An explicit formula for $S*C$, where S is an element from this generating set, can be tediously computed and, from which, the shuffling of the mass can be seen.

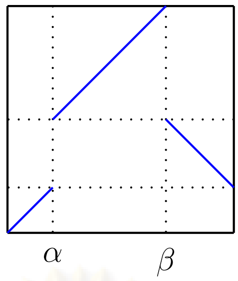

Figure 1.4: the support of a generator where the third stripe is flipped

Let $S_{\alpha,\beta}$ be the shuffle of M whose support is as in Figure 1.4. Then,

$$(S_{\alpha,\beta} * C)(u,v) = \begin{cases} C(u,v) & \text{if } \ 0 \le u \le \alpha \\ C(\alpha,v) + C(u+1-\beta,v) - C(1+\alpha-\beta,v) & \text{if } \ \alpha \le u \le \beta \\ C(\alpha,v) + v - C(1+\alpha-u,v) & \text{if } \ \beta \le u \le 1. \end{cases}$$

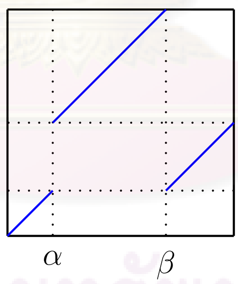

Figure 1.5: the support of a generator where the third stripe is straight

Let $S_{\alpha,\beta}$ be the shuffle of M whose support is as in Figure 1.5. Then,

$$(S_{\alpha,\beta} * C)(u,v) = \begin{cases} C(u,v) & \text{if } \ 0 \le u \le \alpha \le 1 \\ C(\alpha,v) + C(u+1-\beta,v) - C(1+\alpha-\beta,v) & \text{if } \ \alpha \le u \le \beta \\ C(u+\alpha-\beta,v) + v - C(1+\alpha-\beta,v) & \text{if } \ \beta \le u \le 1. \end{cases}$$

Remark 1.11 is very useful when we want to determine the support of a shuffle of a copula. We will use this technique in Example 2.16.

We are also interested in a generalization of the $*$-product. The motivation behind this generalization comes from a research on compatibility of copulas. Copulas C_{12}, C_{13} and C_{23} are said to be compatible if there exists a trivariate copula $\tilde{C}$ such that

$$C_{12}(u,v) = \tilde{C}(u,v,1),$$
$$C_{13}(u,w) = \tilde{C}(u,1,w),$$
$$C_{23}(v,w) = \tilde{C}(1,v,w).$$

Given copulas A and B, the class $\mathcal{C}(A,B)$ is the set of all copulas that are compatible with A and B. To characterize these classes, Durante et al. [3] introduced a binary operation on $\mathfrak{C}$ defined, for each family of copulas $\mathbf{C} = \{C_t\}_{t\in[0,1]}$, by

$$(A *_{\mathbf{C}} B)(u,v) = \int_0^1 C_t(\partial_2 A(u,t), \partial_1 B(t,v))\, dt.$$

They called this operation $\mathbf{C}$-*product*, but in this thesis, we will call it $*_{\mathbf{C}}$ *product*. But it is questionable whether the integrand is Lebesgue measurable for all families of copulas. We will give a detailed discussion on this in Chapter 3.

Example 1.12 ([4], p. 237)**.** For every copula $B \in \mathfrak{C}$ and every family of copulas $\mathbf{C} = \{C_t\}_{t\in[0,1]}$ such that the $*_{\mathbf{C}}$ is well-defined, we have

$$B *_{\mathbf{C}} M = B = M *_{\mathbf{C}} B,$$
$$(B *_{\mathbf{C}} W)(u,v) = u - B(u, 1-v),$$
$$(W *_{\mathbf{C}} B)(u,v) = v - B(1-u, v).$$

Now, we move on to the next definition. In this thesis, we are interested in the modified Sobolev norm introduced by Siburg and Stoimenov [11].

Definition 1.13. For any $A \in \operatorname{span} \mathfrak{C}$, define a norm of A by

$$\|A\| = \left(\int_0^1 \int_0^1 (\partial_1 A(x,y))^2 + (\partial_2 A(x,y))^2 \, dxdy \right)^{1/2}.$$

With a slight abuse of notation, the restriction of $\|\cdot\|$ to $\mathfrak{C}$ is called the *Sobolev norm for copulas.*

The following are some useful properties of the $*$-product and the norm (see, e.g., [1, 11, 12]).

Proposition 1.14. *Let $A, B, C \in \mathfrak{C}$. Then the following statements hold.*

1. $A * (B * C) = (A * B) * C$.
2. $(B * C)^T = C^T * B^T$.
3. $\|C^T\| = \|C\|$.
4. $\|C - \Pi\|^2 = \|C\|^2 - \frac{2}{3}$.
5. $\|C\| = 1$ *if and only if C is invertible.*
6. *The $*$-product is jointly continuous with respect to the norm.*

Definition 1.15. Let X and Y be random variables. Then X is said to be *completely dependent on Y* if there exists a Borel measurable transformation h such that $X = h(Y)$ with probability one. Moreover, X and Y are said to be *mutually completely dependent* if X is completely dependent on Y and Y is completely dependent on X, i.e., there exists a Borel measurable bijective transformation h such that $Y = h(X)$ with probability one.

The following theorem gives some stochastic intepretations of the Sobolev norm for copulas.

Theorem 1.16 ([12], Theorem 4.3)**.** *Let X and Y be continuous random variables with copula C. Then $\|C\|^2 \in [2/3, 1]$. Moreover, the following assertions hold:*

1. $\|C\|^2 = 2/3$ *if and only if X and Y are independent.*

2. $\|C\|^2 \in (5/6, 1]$ *if X is completely dependent on Y or Y is completely dependent on X.*

3. $\|C\|^2 = 1$ *if and only if X and Y are mutually completely dependent.*

We end this section by exploring a relationship between the $*$-product and the Sobolev norm for copulas. In the following example, we observe the $*$-product of a shuffle of M with another copula.

Example 1.17. We consider a family called the Farlie-Gumbel-Morgenstern (FGM) copulas: $C_\theta(u, v) = uv + uv\theta(1 - u)(1 - v)$, $\theta \in [-1, 1]$. Let $S_{1/2}$ be the straight shuffle of M at $1/2$, which is an invertible copula. Then it can be computed, though tediously, that

$$\|S_{1/2} * C_\theta\|^2 = \|C_\theta\|^2 - \tfrac{5}{144}\theta^2 < \|C_\theta\|^2 \text{ for all } \theta \neq 0.$$

Hence, a map $A \mapsto U * A * V$ where $U, V \in \operatorname{Inv} \mathfrak{C}$ is, in general, not an isometry with respect to the Sobolev norm for copulas.

CHAPTER II

A COMPATIBLE NORM

2.1 Shuffling maps

In this section, we introduce the definition of a shuffling map on the set span $\mathfrak{C}$. Then we survey some of its stochastic properties.

Definition 2.1. Let $U, V \in \operatorname{Inv}\mathfrak{C}$, the set of invertible copulas. A shuffling map $S_{U,V}$ is a map on span $\mathfrak{C}$ defined by

$$S_{U,V}(A) = U * A * V.$$

The motivation behind the word "shuffling" comes from the fact that a shuffling image of a copula is a two-sided generalized shuffle of the copula, which was introduced in Chapter 1.

Proposition 2.2 ([1], p. 610)**.** *If Z and Y are conditionally independent given X, then $C_{Z,Y} = C_{Z,X} * C_{X,Y}$.*

Proposition 2.3. *Let $h\colon \mathbb{R} \to \mathbb{R}$ be Borel measurable. Then, for any random variables X, Y, we have $h(X)$ and Y are conditionally independent given X.*

Proof. Observe that $h(X) \in \sigma(X)$, the σ-algebra generated by X. Hence, by properties of conditional expectations,

$$\begin{aligned} E(I_{h(X)\le a}|X)(\omega) \cdot E(I_{Y\le b}|X)(\omega) &= I_{h(X)\le a}(\omega) \cdot E(I_{Y\le b}|X)(\omega) \\ &= E(I_{h(X)\le a} \cdot I_{Y\le b}|X)(\omega) \end{aligned}$$

for all $\omega \in \Omega$. This completes the proof. □

Corollary 2.4. *Let $f, g\colon \mathbb{R} \to \mathbb{R}$ be Borel measurable functions. Then*

$$C_{f(X),X} * C_{X,Y} * C_{Y,g(Y)} = C_{f(X),g(Y)}$$

for all random variables X, Y.

Proof. Since f and g are Borel measurable, by Propositions 2.2 and 2.3, we have

$$C_{f(X),Y} = C_{f(X),X} * C_{X,Y} \quad \text{and} \tag{2.1}$$

$$C_{g(Y),X} = C_{g(Y),Y} * C_{Y,X} \tag{2.2}$$

for all random variables X, Y. Transpose both sides of (2.2), we obtain $C_{X,g(Y)} = C_{X,Y} * C_{Y,g(Y)}$. Then, we have

$$\begin{aligned} C_{f(X),g(Y)} &= C_{f(X),X} * C_{X,g(Y)} && (2.3)\\ &= C_{f(X),X} * C_{X,Y} * C_{Y,g(Y)}. && (2.4) \end{aligned}$$

□

We are now ready to derive stochastic properties of shuffling maps.

Lemma 2.5. *Let X, Y be continuous random variables and $U, V \in \operatorname{Inv}\mathfrak{C}$. Then the following statements hold:*

1. *X and Y are independent if and only if $S_{U,V}(C_{X,Y}) = \Pi$.*
2. *X is completely dependent on Y or vice versa if and only if $S_{U,V}(C_{X,Y})$ is a complete dependence copula.*
3. *X and Y are mutually completely dependent if and only if $S_{U,V}(C_{X,Y})$ is a mutual complete dependence copula.*

Proof. We only need to prove the implications as the converses automatically follow because the inverse of a shuffling map is still a shuffling map.

1. The result is clear because the copula $C_{X,Y} = \Pi$ if and only if X, Y are independent.

2. With out loss of generality, assume that Y is completely dependent on X. Then there exists a Borel measurable transformation h such that $Y = h(X)$ with probability one. Consider a shuffling map $S_{U,V}$. There exist Borel measurable bijective transformations f, g such that $U = C_{f(X),X}$ and $V = C_{Y,g(Y)}$. By Corollary 2.4, we have

$$S_{U,V}(C_{X,Y}) = C_{f(X),X} * C_{X,Y} * C_{Y,g(Y)} = C_{f(X),g(Y)}.$$

Thus, it suffices to show that $g(Y)$ is completely dependent on $f(X)$. From $Y = h(X)$ with probability one, $g(Y) = (g \circ h)(X) = (g \circ h \circ f^{-1})(f(X))$ with probability one. It is left to show that f^{-1} is Borel measurable. This is true because of Lusin-Souslin Theorem (see, e.g., [6], Corollary 15.2) which states that a Borel measurable injective image of a Borel set is a Borel set.

3. The proof is similar to the proof above except that the functions h and $g \circ h \circ f^{-1}$ are now Borel measurable bijective transformations instead of Borel measurable transformations. □

Corollary 2.4 implies that a shuffling image of a copula $C_{X,Y}$ is a copula of transformed random variables $C_{f(X),g(Y)}$ for some Borel measurable bijective transformations f and g. Together with the above lemma, we obtain the following theorem.

Theorem 2.6. *Let X and Y be continuous random variables. Let f and g be any Borel measurable bijective transformations of the random variables X and Y, respectively. Then X and Y are independent, completely dependent or mutually completely dependent if and only if $f(X)$ and $g(Y)$ are independent, completely dependent or mutually completely dependent, respectively.*

The above theorem suggests that shuffling maps preserve stochastic properties of copulas. In the next section, we contruct a norm which, in some sense, also preserves stochastic properties of copulas.

2.2 The $*$-norm

Our main purpose is to construct a norm under which shuffling maps are isometries and then derive its properties.

Definition 2.7. Denoted by $\|\cdot\|$ the Sobolev norm for copulas. Define a map $\|\cdot\|_* : \operatorname{span}\mathfrak{C} \to [0,\infty)$, by

$$\|A\|_* = \sup_{U,V\in \operatorname{Inv}\mathfrak{C}} \|U * A * V\|.$$

It can be easily checked that $\|\cdot\|_*$ is a norm on $\operatorname{span}\mathfrak{C}$. We call $\|\cdot\|_*$ the $*$-norm. Moreover, from the definition, it is clear that $\|A\| \le \|A\|_*$ for all $A \in \operatorname{span}\mathfrak{C}$.

Lemma 2.8. *Let $U, V \in \mathfrak{C}$. If $\|U\| = 1$, then $\|V\| = 1$ if and only if $\|U * V\| = 1$. Similarly, if $\|V\| = 1$, then $\|U\| = 1$ if and only if $\|U * V\| = 1$.*

Proof. It suffices to prove only the first statement as the second statement can be proved similarly.

Let $U, V \in \mathfrak{C}$ be such that $\|U\| = 1$ and $\|U * V\| = 1$. Then U, V are invertible. We know the set of shuffles of M is dense in $\operatorname{Inv}\mathfrak{C}$ with respect to the Sobolev norm. Then, with respect to the Sobolev norm, there exist S_n, T_n shuffles of

M such that $S_n \to U$ and $T_n \to V$. Hence, with respect to the Sobolev norm, $S_n * T_n \to U * V$ by the joint continuity of the $*$-product. But a product of shuffles of M is still a shuffle of M, which is invertible. Hence, $\|U * V\| = 1$.

Let U and $U * V$ be copulas of Sobolev norm 1. Since $\|U^T\| = \|U\| = 1$, then, by the previous result, we have $\|V\| = \|U^T * (U * V)\| = 1$. □

We move on to deriving properties of the $*$-norm. The following proposition summarizes the results used to contruct and derive properties of the measure of dependence.

Proposition 2.9. *Let $A \in \operatorname{span} \mathfrak{C}$ and $C \in \mathfrak{C}$. Then the following statements hold.*

1. $\|C\|_*^2 = \frac{2}{3}$ *if and only if* $C = \Pi$.
2. $\|C\|_* = 1$ *if* $\|C\| = 1$.
3. $\|C - \Pi\|_*^2 = \|C\|_*^2 - \frac{2}{3}$.
4. *Transposition map is an isometry with respect to the $*$-norm.*

Proof. 1. Since $U * \Pi * V = \Pi$ for all $U, V \in \operatorname{Inv} \mathfrak{C}$, the result is clear.

2. If $\|C\| = 1$, then $1 = \|C\| \leq \|C\|_* \leq 1$.

3. From property 4 of Proposition 1.14, we have

$$\|U * (C - \Pi) * V\|^2 = \|U * C * V - \Pi\|^2 = \|U * C * V\|^2 - \tfrac{2}{3}$$

for all $U, V \in \operatorname{Inv} \mathfrak{C}$. The result follows by taking supremum over $U, V \in \operatorname{Inv} \mathfrak{C}$ on both sides.

4. Let $A \in \operatorname{span}\mathfrak{C}$. We know $\|U^T\| = \|U\|$ for all $U \in \mathfrak{C}$. In particular, $U^T \in \operatorname{Inv} C$ if and only if $U \in \operatorname{Inv} C$. Hence,

$$\begin{aligned}
\|A^T\|_* &= \sup_{U,V \in \operatorname{Inv}\mathfrak{C}} \|U * A^T * V\| \\
&= \sup_{U,V \in \operatorname{Inv}\mathfrak{C}} \|V^T * A * U^T\| \\
&= \sup_{U^T,V^T \in \operatorname{Inv}\mathfrak{C}} \|V^T * A * U^T\| \\
&= \sup_{U,V \in \operatorname{Inv}\mathfrak{C}} \|U * A * V\| \\
&= \|A\|_*.
\end{aligned}$$

□

Theorem 2.10. *Shuffling maps are isometries with respect to the $*$-norm.*

Proof. Let $A \in \operatorname{span}\mathfrak{C}$ and $B, C \in \operatorname{Inv}\mathfrak{C}$. Then, for any $U \in \mathfrak{C}$, $\|U * B\| = 1$ if and only if $\|U\| = 1$ by Lemma 2.8. In other words, $U * B \in \operatorname{Inv}\mathfrak{C}$ if and only if $U \in \operatorname{Inv}\mathfrak{C}$. Similarly, for any $V \in \mathfrak{C}$, $C * V \in \operatorname{Inv}\mathfrak{C}$ if and only if $V \in \operatorname{Inv}\mathfrak{C}$. Hence,

$$\begin{aligned}
\|B * A * C\|_* &= \sup_{U,V \in \operatorname{Inv}\mathfrak{C}} \|(U * B) * A * (C * V)\| \\
&= \sup_{U*B,C*V \in \operatorname{Inv}\mathfrak{C}} \|(U * B) * A * (C * V)\| \\
&= \sup_{U,V \in \operatorname{Inv}\mathfrak{C}} \|U * A * V\| \\
&= \|A\|_*.
\end{aligned}$$

□

Here, we give two examples: the first example suggests that the Sobolev norm and the $*$-norm are distinct and the other gives a class of copulas on which the Sobolev norm and the $*$-norm are equal.

Example 2.11. From the setup of Example 1.17, let $A_\theta = S_{1/2} * C_\theta$. Then, we have $S_{1/2} * A_\theta = S_{1/2} * S_{1/2} * C_\theta = C_\theta$. Also from Example 1.17, we have that $\|A_\theta\| < \|C_\theta\|$ for any $\theta \neq 0$. Then

$$\|A_\theta\|_* \geq \|S_{1/2} * A_\theta\| = \|C_\theta\| > \|A_\theta\|.$$

Hence, the two norms are distinct.

Example 2.12. Consider the family of convex sums of an invertible copula and the product copula $\{\alpha A + (1-\alpha)\Pi\}_{\alpha \in [0,1]}$. Then

$$\begin{aligned}
\|\alpha A + (1-\alpha)\Pi\|_*^2 &= \|\alpha A + (1-\alpha)\Pi - \Pi\|_*^2 + \frac{2}{3} \\
&= \|\alpha A - \alpha\Pi\|_*^2 + \frac{2}{3} \\
&= \alpha^2\|A - \Pi\|_*^2 + \frac{2}{3} \\
&= \alpha^2(\|A\|_*^2 - \frac{2}{3})^2 + \frac{2}{3} \\
&= \alpha^2(\|A\|^2 - \frac{2}{3})^2 + \frac{2}{3} \\
&= \alpha^2\|A - \Pi\|^2 + \frac{2}{3} \\
&= \|\alpha A - \alpha\Pi\|^2 + \frac{2}{3} \\
&= \|\alpha A + (1-\alpha)\Pi\|^2.
\end{aligned}$$

Hence, the Sobolev norm and the $*$-norm coincide on the family of convex sums of an invertible copula and the product copula.

Remark 2.13. At first, we thought that the set $\operatorname{Inv}\mathfrak{C}$ is compact with respect to the Sobolev norm for copulas, but that is not the case. If this were true, we would have obtained that, for any $A \in \operatorname{span}\mathfrak{C}$, there exist $U, V \in \operatorname{Inv}\mathfrak{C}$ such that $\|A\|_* = \|U * A * V\|$. Consequently, we would have the converse of the second statement in Proposition 2.9. However, this is false as a counterexample is given in Example 2.16. Before that, let us discuss why $\operatorname{Inv}\mathfrak{C}$ is not compact with respect to the Sobolev norm for copulas.

Definition 2.14. For any $k \in \mathbb{N}$, $p \in [1, \infty]$ and $\Omega \subseteq \mathbb{R}^n$, the Sobolev space $W^{k,p}(\Omega)$ is defined to be the set of all functions $f \in L^p(\Omega)$ such that for every multi-index α with $|\alpha| \leq k$, the weak partial derivative $D^\alpha u \in L^p(\Omega)$.

Proposition 2.15 ([2], pp. 426, 436)**.** *For any $p \in [1, \infty]$, let $\| \cdot \|_{1,p}$ be the classical Sobolev norm defined on the Sobolev space $W^{1,p}([0,1]^2)$ and let $| \cdot |_{1,p}$ be the corresponding seminorm defined by*

$$|f|_{1,p} = \left(\int_0^1 \int_0^1 (|\partial_1 f(x,y)|^p + |\partial_2 f(x,y)|^p) \; dxdy \right)^{1/p}.$$

Then, the following statements hold.

1. *$| \cdot |_{1,p}$ restricted to* $\operatorname{span} \mathfrak{C}$ *is a norm.*
2. *On* $\operatorname{span} \mathfrak{C}$*, $| \cdot |_{1,p}$ dominates the uniform norm for all $p \in (2, \infty]$.*
3. *The set* $\operatorname{Inv} \mathfrak{C}$ *is not compact with respect to any norm which dominates or is dominated by the uniform norm on* $\operatorname{span} \mathfrak{C}$*.*
4. *The topologies generated by $| \cdot |_{1,p}$ coincide on $\mathfrak{C}$ for all $p \in [1, \infty]$.*

From properties 2 and 3 of the above proposition, $\operatorname{Inv} \mathfrak{C}$ is not compact with respect to any $| \cdot |_{1,p}$ where $p \in (2, \infty]$. Together with property 4 of the same proposition, we can conclude that $\operatorname{Inv} \mathfrak{C}$ is not compact with respect to any $| \cdot |_{1,p}$ where $p \in [1, \infty]$. In particular, it is not compact with respect to the Sobolev norm for copulas. Furthermore, $\operatorname{Inv} \mathfrak{C}$ is complete with respect to the Sobolev norm for copulas because it is a closed subset of the set of copulas, which is complete (for a proof see [2], Theorem 4.5). As a consequence, $\operatorname{Inv} \mathfrak{C}$ is not totally bounded with respect to the Sobolev norm for copulas since it is a complete metric space which is not compact.

In the following example, we give a copula $C_0 \in \mathfrak{C}$ such that $\|C_0\| \neq 1$ but $\|C_0\|_* = 1$. To show that $\|C_0\|_* = 1$, we construct a sequence of invertible copulas $U_n \in \operatorname{Inv}\mathfrak{C}$ such that $\|U_n * C_0\| \to 1$.

Example 2.16. Let C_0 be the copula supported on the straight line joining the points $(0,0)$ and $(1/2,1)$ and the straight line joining the points $(1/2,0)$ and $(1,1)$. It is known that C_0 is not invertible; hence, $\|C_0\| < 1$. Consider a partition of $[0,1]^2$ into 2^{n+1} equal vertical stripes where $n \in \mathbb{N}$. Let S_n be the shuffle of M which switches, for all $j \equiv 2 \mod 4$, the supports of the j-th and $(j+1)$-th stripes of M. For a better understanding of this construction, see the figure below.

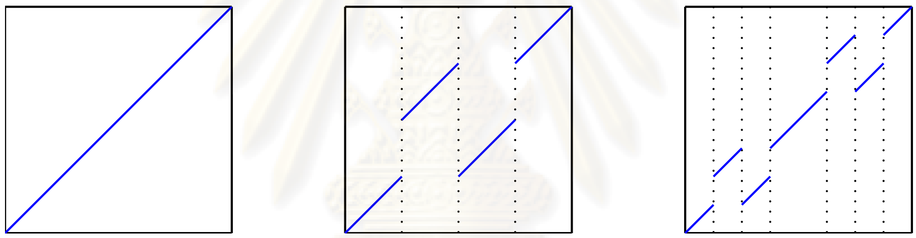

Figure 2.1: the supports of M, S_1 and S_2

Construct, recursively, $C_n = S_n * C_{n-1}$ for all $n \in \mathbb{N}$. According to Remark 1.11, the support of C_{n-1} is shuffled according to the shuffling of S_n, i.e., for all $j \equiv 2 \mod 4$, the supports of the j-th and $(j+1)$-th stripes of C_{n-1} are switched. Then the support of C_n lies entirely in the diagonal 2^n-squares. For examples, see the figure below.

Therefore, by applying the technique demonstrated in Example 1.5, it can be shown that the copula C_n and M coincide on the area outside the diagonal 2^n-squares. But the union of the diagonal 2^n-squares is a descending chain, the intersection of which is the diagonal of $[0,1]^2$ joining the points $(0,0)$ and $(1,1)$. This implies that, for $i = 1,2$, we have $\partial_i C_n(x,y) \to \partial_i M(x,y)$ pointwise for all

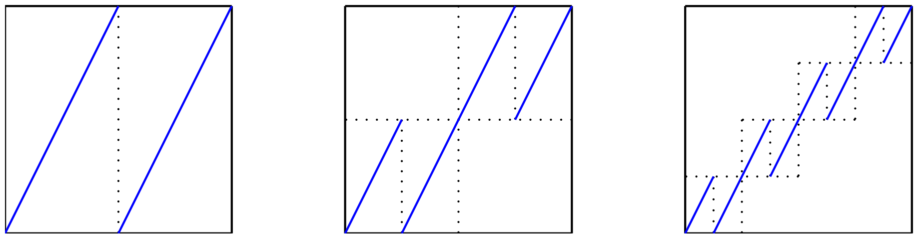

Figure 2.2: the supports of C_0, C_1 and C_2

$x \neq y$. Note that, for fixed $x, y \in [0,1]$ such that $x \neq y$, there exists $N \in \mathbb{N}$ large enough so that the point (x, y) is outside the 2^N-squares. Hence, for each $i = 1, 2$, $\partial_i C_n(x, y)$ exists for all $n \geq N$. Thus we have $|\nabla C_n(x, y)|^2 \to |\nabla M(x, y)|^2$ a.e. $(x, y) \in [0,1]^2$ and they are bounded by 4 which is integrable on $[0,1]^2$. Hence, by Dominated Convergence Theorem, we have $\|C_n\| \to \|M\| = 1$.

To sum up, $C_n = S_n * C_{n-1} = \cdots = (S_n * S_{n-1} * \cdots * S_1) * C_0$ is a product of a shuffle of M and the copula C_0 for all $n \in \mathbb{N}$. So we have a copula C_0 and sequences of invertible copulas $\{U_n\}$ and $\{V_n\}$ where $U_n = S_n * S_{n-1} * \cdots * S_1$ and $V_n = M$ such that $\|C_0\| < 1$ but $\|U_n * C_0 * V_n\| \to 1$, which implies $\|C_0\|_* = 1$. Therefore, for $C \in \mathfrak{C}$, $\|C\|_* = 1$ does not imply $\|C\| = 1$.

2.3 The measure ω_*

First, let's recall the definition and properties of the measure of mutual complete dependence ω introduced by Siburg and Stoimenov [12]. Let X and Y be continuous random variables with copula C. The measure ω is defined by $\omega(X, Y) = \sqrt{3}\|C - \Pi\|$, which can be viewed as the normalized Sobolev distance between the copula C and the independence copula. The following theorem summarizes properties of the measure.

Theorem 2.17 ([12], Theorem 5.3). *Let X and Y be continuous random variables with copula C. Then the measure $\omega(X,Y)$ has the following properties:*

1. $\omega(X,Y) = \omega(Y,X)$.

2. $0 \le \omega(X,Y) \le 1$.

3. $\omega(X,Y) = 0$ *if and only if X and Y are independent.*

4. $\omega(X,Y) = 1$ *if and only if X and Y are mutually completely dependent.*

5. $\omega(X,Y) \in (\sqrt{2}/2, 1]$ *if Y is completely dependent on X or X is completely dependent on Y.*

6. *If f and g are monotone transformations, then* $\omega(f(X), g(Y)) = \omega(X,Y)$.

7. *If $\{(X_n, Y_n)\}_{n\in\mathbb{N}}$ is a sequence of pairs of continuous random variables with copulas $\{C_n\}_{n\in\mathbb{N}}$ and if $\lim_{n\to\infty} \|C_n - C\| = 0$, then we have* $\lim_{n\to\infty} \omega(X_n, Y_n) = \omega(X,Y)$.

Analogous to the measure ω, we define a new measure of dependence using the $*$-norm as follows. Let X and Y be continuous random variables with copula C. Define $\omega_*(X,Y) = \sqrt{3}\|C - \Pi\|_*$. Moreover, from property 3 of Proposition 2.9, we have $\omega_*(X,Y) = (3\|C\|_*^2 - 2)^{1/2}$.

Because the properties of the $*$-norm are for the most part analogous to those of the Sobolev norm, the properties of ω_* are consequently analogous to those of ω's except for properties 4 and 6 in Theorem 2.17. The measure ω_*, unlike the measure ω, is not a measure of mutual complete dependence since there exists, according to Example 2.16, a pair of continuous random variables which are not mutually completely dependent but their copula has $*$-norm one. This is the downfall of our measure compared to the measure ω. Nevertheless, for the measure ω_*, we can weaken the assumptions on the transformations f and g in property 6 in Theorem 2.17. In order to do this, we use the fact that the shuffling maps are isometries with respect to the $*$-norm. The following theorem summarizes properties of the measure ω_*.

Theorem 2.18. *Let X and Y be continuous random variables with copula C. Then $\omega_*(X,Y)$ has the following properties:*

1. $\omega_*(X,Y) = \omega_*(Y,X)$.

2. $0 \leq \omega_*(X,Y) \leq 1$.

3. $\omega_*(X,Y) = 0$ *if and only if X and Y are independent.*

4. $\omega_*(X,Y) = 1$ *if X and Y are mutually completely dependent.*

5. $\omega_*(X,Y) \in (\sqrt{2}/2, 1]$ *if Y is completely dependent on X or X is completely dependent on Y.*

6. *If f and g are Borel measurable bijective transformations, then we have* $\omega_*(f(X), g(Y)) = \omega_*(X,Y)$.

7. *If $\{(X_n, Y_n)\}_{n\in\mathbb{N}}$ is a sequence of pairs of continuous random variables with copulas $\{C_n\}_{n\in\mathbb{N}}$ and if $\lim_{n\to\infty} \|C_n - C\|_* = 0$, then we have $\lim_{n\to\infty} \omega_*(X_n, Y_n) = \omega_*(X,Y)$.*

Proof. Let X and Y be continuous random variables.

1. This follows from the fact that $\|C_{X,Y}\|_* = \|C_{Y,X}\|_*$ since, by Proposition 2.9, the transposition map is an isometry with respect to the $*$-norm.

2. By Theorem 1.16, $\|C_{X,Y}\|^2 \in [2/3, 1]$. Since $\|A\|_* \geq \|A\|$ for all $A \in \operatorname{span}\mathfrak{C}$, we have $\|C_{X,Y}\|_*^2 \in [2/3, 1]$. Hence, $0 \leq \omega_*(X,Y) \leq 1$.

3. By Proposition 2.9, X and Y are independent if and only if $C_{X,Y} = \Pi$ if and only if $\|C_{X,Y}\|_*^2 = 2/3$. Therefore, X and Y are independent if and only if $\omega_*(X,Y) = 0$.

4. If X and Y are mutually completely dependent, then $\|C_{X,Y}\|_* = 1$. Therefore, $\omega_*(X,Y) = 1$.

5. Let Y be completely dependent on X or X be completely dependent on Y. Then $\|C_{X,Y}\|^2 \in (5/6, 1]$ by Theorem 1.16. Since $\|A\|_* \geq \|A\|$ for all $A \in \operatorname{span}\mathfrak{C}$, we have $\|C_{X,Y}\|_*^2 \in (5/6, 1]$. Hence, $\omega_*(X,Y) \in (\sqrt{2}/2, 1]$.

6. Let f, g be Borel measurable bijective transformations. Then, X and $f(X)$ are mutually completely dependent, and so are Y and $g(Y)$. Thus $\|C_{f(X),X}\| = 1$ and $\|C_{Y,g(Y)}\| = 1$ by Theorem 1.16. Therefore, the copulas $C_{f(X),X}$ and $C_{Y,g(Y)}$ are invertible by property 5 in Proposition 1.14. Hence

$$\begin{aligned}
\omega_*(f(X), g(Y)) &= \sqrt{3}\|C_{f(X),g(Y)} - \Pi\|_* \\
&= \sqrt{3}\|C_{f(X),X} * (C_{X,Y} - \Pi) * C_{Y,g(Y)}\|_* \\
&= \sqrt{3}\|C_{X,Y} - \Pi\|_* \\
&= \omega_*(X,Y).
\end{aligned}$$

7. From the definition of ω_*, observe that

$$\begin{aligned}
|\omega_*(X_n, Y_n) - \omega_*(X,Y)| &= \sqrt{3}\big|\|C_n - \Pi\|_* - \|C - \Pi\|_*\big| \\
&\leq \sqrt{3}\|C_n - C\|_*.
\end{aligned}$$

□

We end this chapter with the list of open problems we encountered during our work on this part of thesis.

1. Besides the transposition and shuffling maps and their compositions, are there any other maps which are isometries with respect to the $*$-norm?

2. What are necessary and sufficient conditions on a copula C with $\|C\|_* = 1$?

3. What is the set of copulas on which the Sobolev norm and the $*$-norm coincide?

4. Is the $*$-product jointly continuous with respect to the $*$-norm?

5. What are probabilistic interpretations of shuffling maps?

CHAPTER III

A GENERALIZED $*$-PRODUCT

3.1 Measurability of the integrand

In this section, we introduce various sets of conditions on the family of copulas so that the $*_{\mathbf{C}}$ product is well-defined.

Example 3.1. Let P be a Lebesgue nonmeasurable subset of $[0,1]$. Consider the family $\mathbf{C} = \{C_t\}_{t\in[0,1]}$ where

$$C_t = \begin{cases} M & \text{if } \ t \in P \\ W & \text{if } \ t \notin P. \end{cases}$$

Then we can see that $C_t(\partial_2 A(x,t), \partial_1 B(t,y))$ is not Lebesgue measurable in the variable t for some $A, B \in \mathfrak{C}$ and $x, y \in [0,1]$, e.g., $A, B = \Pi$ and any $x, y \in (0,1)$ so that $M(x,y) > W(x,y)$.

From the above example, it is evident that the $*_{\mathbf{C}}$ product is not always well-defined since the integrand may not be Lebesgue measurable. One way to solve this measurability problem is to restrict our attention to smaller classes of families of copulas. We give two sets of conditions such that $C_t(\partial_2 A(x,t), \partial_1 B(t,y))$ is a Lebesgue measurable function in the variable t.

The first set of conditions is given in the following theorem. Practically, almost all families of copulas we encounter satify this set of conditions.

Theorem 3.2. *Let* $\mathbf{C} = \{C_t\}_{t\in[0,1]}$ *be a family of copulas which satisfies*

1. $\mathbf{C}$ *consists of countably many distinct copulas and*

2. *for each* $A \in \mathfrak{C}$*, the set* $\{t \in [0,1] \colon C_t = A\}$ *is Borel measurable.*

Then, for all $x, y \in [0,1]$*,* $C_t(\partial_2 A(x,t), \partial_1 B(t,y))$ *is Lebesgue measurable in the variable* t*.*

Proof. Let $\mathbf{C} = \{C_t\}_{t\in[0,1]}$ be a family of copulas satisfying the two conditions. Since there are countably many distinct copulas. Let $E = \{C_1, C_2, \dots\}$ be an enumeration of the distinct copulas in the family.

For each $C_n \in E$, let $T_n = \{t \in [0,1] \colon C_t = C_n\}$. Observe that $\{T_n\}_{n=1}^{\infty}$ is a partition of $[0,1]$ into measurable sets. Then, we write $C_t(\partial_2 A(x,t), \partial_1 B(t,y))$ as

$$\sum_{n=1}^{\infty} \chi_{T_n}(t) C_n(\partial_2 A(x,t), \partial_1 B(t,y)),$$

which is a countable sum of Lebesgue measurable functions; hence, it is Lebesgue measurable. □

Observe that the proof of the above theorem works perfectly fine if we replace Borel measurability by Lebesgue measurability.

Theorem 3.3. *If the map* $(t,x,y) \mapsto C_t(x,y)$ *is Borel measurable, then for all* $x, y \in [0,1]$ *and for all* $A, B \in \mathfrak{C}$*,* $C_t(\partial_2 A(x,t), \partial_1 B(t,y))$ *is Lebesgue measurable in the variable* t*.*

Proof. For any fixed $x, y \in [0,1]$, the map $t \mapsto (t, \partial_2 A(x,t), \partial_1 B(t,y))$ is Lebesgue measurable since each component function is Lebesgue measurable. Then, being the composition of a Lebesgue measurable map $t \mapsto (t, \partial_2 A(x,t), \partial_1 B(t,y))$ and a Borel measurable map $(t,x,y) \mapsto C_t(x,y)$, the map $t \mapsto C_t(\partial_2 A(x,t), \partial_1 B(t,y))$ is Lebesgue measurable. □

Denoted by $\mathcal{M}_c$ the collection of families which satisfy the set of conditions in Theorem 3.2, $\mathcal{M}_u$ the collection of families which satisfy the condition in Theorem 3.3 and $\mathcal{M}$ the collection of families $\{C_t\}_{t\in[0,1]}$ such that, for all $A, B \in \mathfrak{C}$ and $x, y \in [0,1]$, the function $C_t(\partial_2 A(x,t), \partial_1 B(t,y))$ is Lebesgue measurable in the variable t. We have just shown that $\mathcal{M}_c$ and $\mathcal{M}_u$ are subcollections of $\mathcal{M}$. Let's remark that, in practice, it is not easy to determine whether a family is in $\mathcal{M}$. This is the sole reason we introduce the collections $\mathcal{M}_c$ and $\mathcal{M}_u$.

Lemma 3.4. *If a family of copulas satifies the set of conditions in Theorem 3.2, then it also satisfies the condition in Theorem 3.3. In other words,* $\mathcal{M}_c \subseteq \mathcal{M}_u$.

Proof. Let $\mathbf{C} = \{C_t\}_{t\in[0,1]} \in \mathcal{M}_c$. For each $C_n \in \mathbf{C}$, let $T_n = \{t \in [0,1] \colon C_t = C_n\}$. Then we can write

$$C_t(x,y) = \sum_{n=1}^{\infty} \chi_{T_n}(t) C_n(x,y).$$

Now, for any $a \in [0,1]$, the inverse image of the interval $[0,a]$ under the map $(t,x,y) \mapsto C_t(x,y)$ is equal to $\bigcup_{n=1}^{\infty} T_n \times C_n^{-1}([0,a])$. Observe that T_n and $C_n^{-1}([0,a])$ are Borel measurable. Hence, the inverse image of the interval $[0,a]$ under the map $(t,x,y) \mapsto C_t(x,y)$ is Borel measurable. □

The following proposition helps us in dealing with families which behave well outside a set of Lebesgue measure zero.

Proposition 3.5. *Let* $\mathbf{C} \in \mathcal{M}$. *If* $\mathbf{D}$ *is another family of copulas such that* $D_t = C_t$ *a.e.* $t \in [0,1]$, *then* $\mathbf{D} \in \mathcal{M}$ *and the products* $*_{\mathbf{C}}$ *and* $*_{\mathbf{D}}$ *are identical. We say that the family* $\mathbf{D}$ *is* $*$*-equivalent to the family* $\mathbf{C}$.

Proof. The result easily follows from the fact that if $f = g$ a.e. and f is Lebesgue measurable, then g is also Lebesgue measurable. Moreover, for any Lebesgue measurable set A, $\int_A f \, d\lambda = \int_A g \, d\lambda$ where λ denotes Lebesgue measure. □

3.2 The $*_{\mathbf{C}}$ product

In this section, we properly re-define the $*_{\mathbf{C}}$ product. Then we derive some of its properties.

Definition 3.6. Let $\mathbf{C} \in \mathcal{M}$. The $*_{\mathbf{C}}$ product of copulas A and B is defined by

$$(A *_{\mathbf{C}} B)(x,y) = \int_0^1 C_t(\partial_2 A(x,t), \partial_1 B(t,y))\ dt.$$

Remark 3.7 ([3], Proposition 3.1)**.** For all $\mathbf{C} \in \mathcal{M}$ and for all $A, B \in \mathfrak{C}$, we have $A *_{\mathbf{C}} B \in \mathfrak{C}$.

Lemma 3.8. *Let $\mathbf{C} \in \mathcal{M}$ and $A, B \in \mathfrak{C}$. If A is right invertible or B is left invertible with respect to the $*$-product, then $A *_{\mathbf{C}} B = A * B$.*

Proof. It suffices to prove only the first statement as the second is analogous.

Let A be a right invertible copula. Then $\partial_2 A(x,y) \in \{0,1\}$ almost everywhere. Let Z be the set $\{(x,y) \colon \partial_2 A(x,y) = 1\}$. Compute

$$\begin{aligned}
(A *_{\mathbf{C}} B)(x,y) &= \int_0^1 C_t(\partial_2 A(x,t), \partial_1 B(t,y))\ dt \\
&= \int_Z C_t(1, \partial_1 B(t,y))\ dt \\
&= \int_Z \partial_1 B(t,y)\ dt \\
&= \int_0^1 \partial_2 A(x,t) \partial_1 B(t,y)\ dt \\
&= (A * B)(x,y).
\end{aligned}$$

$\square$

Theoretically, we often encounter the $*_{\mathbf{C}}$ of copulas A and B where one of them is invertible, so the above lemma is very useful.

Lemma 3.9. *$\{\Pi *_{\mathbf{C}} \Pi \colon \mathbf{C} \in \mathcal{M}\} = \mathfrak{C}$.*

Proof. For any copula $C \in \mathfrak{C}$, consider the family $\mathbf{C}$ consisting of $C_t = C$ for all $t \in [0,1]$. Then, we have $\mathbf{C} \in \mathcal{M}_c$ and

$$(\Pi *_{\mathbf{C}} \Pi)(x,y) = \int_0^1 C(x,y)\ dt = C(x,y).$$

This completes the proof. □

Theorem 3.10. *If $\mathbf{C}_n, \mathbf{C} \in \mathcal{M}$ such that $C_{n,t}(x,y) \to C_t(x,y)$ pointwise for all $t \in [0,1]$, then $(A *_{\mathbf{C}_n} B)(x,y) \to (A *_{\mathbf{C}} B)(x,y)$ pointwise.*

Proof. Observe that, for a fixed $t \in [0,1]$,

$$C_{n,t}(\partial_2 A(x,t), \partial_1 B(t,y)) \to C_t(\partial_2 A(x,t), \partial_1 B(t,y))$$

pointwise. Moreover, $C_{n,t}(\partial_2 A(x,t), \partial_1 B(t,y)), C_t(\partial_2 A(x,t), \partial_1 B(t,y))$ are bounded by 1 which is Lebesgue integrable on $[0,1]$. By Dominated Convergence Theorem,

$$\int_0^1 C_{n,t}(\partial_2 A(x,t), \partial_1 B(t,y))\ dt \to \int_0^1 C_t(\partial_2 A(x,t), \partial_1 B(t,y))\ dt$$

pointwise. This completes the proof. □

Corollary 3.11. *For any $A, B \in \mathfrak{C}$. If $\mathbf{C}_n, \mathbf{C} \in \mathcal{M}$ such that $C_{n,t} \to C_t$ uniformly for all $t \in [0,1]$, then $A *_{\mathbf{C}_n} B \to A *_{\mathbf{C}} B$ uniformly.*

Example 3.12. Recall that the set of shuffles of M is dense in $\mathfrak{C}$ with respect to the uniform norm. Hence, given a family of copulas $\mathbf{C} = \{C_t\}_{t \in [0,1]}$, we can find families of shuffles of M, $\mathbf{S}_n = \{S_{n,t}\}_{t \in [0,1]}$, such that $A *_{\mathbf{S}_n} B \to A *_{\mathbf{C}} B$ uniformly.

Our motivation for the previous example is the computation of $A *_{\mathbf{C}} B$. One can see that given a family $\mathbf{C} = \{C_t\}_{t \in [0,1]}$, it is not easy to obtain an explicit formula for $A *_{\mathbf{C}} B$. But with the above result, the computation seems more feasible.

3.3 Identity and zero of $(\mathfrak{C}, *_{\mathbf{C}})$

Recall from Chapter 1 that the identity and the zero of $(\mathfrak{C}, *)$ are M and Π, respectively.

Theorem 3.13. *For all $\mathbf{C} \in \mathcal{M}$, the identity for $(\mathfrak{C}, *_{\mathbf{C}})$ exists and is unique. Moreover, it is the Fréchet-Hoeffding upper bound M.*

Proof. Let $\mathbf{C} \in \mathcal{M}$. Since M is invertible, from Lemma 3.8, we have

$$M *_{\mathbf{C}} A = M * A = A = A * M = A *_{\mathbf{C}} M$$

for all $A \in \mathfrak{C}$. For the uniqueness, suppose M' is another identity. Then we have

$$M = M *_{\mathbf{C}} M' = M'.$$

Hence, for all $\mathbf{C} \in \mathcal{M}$, the copula M is the identity for the $*_{\mathbf{C}}$ product. □

Theorem 3.14. *Let $\mathbf{C} \in \mathcal{M}$. A zero for the $*_{\mathbf{C}}$ product, if exists, is unique and is the product copula Π.*

Proof. The uniqueness part is easy. Let U, V be zeroes for the $*_{\mathbf{C}}$ product. Then $U = U *_{\mathbf{C}} V = V$. Now, to see that Π is the zero, if exists, it requires some work.

Let U be the zero for $*_{\mathbf{C}}$. For each S_α, the straight shuffle of M at $\alpha \in [0, 1]$, since S_α is invertible, we have $S_\alpha * U = S_\alpha *_{\mathbf{C}} U = U$.

Recall the formula in Example 1.10:

$$(S_\alpha * C)(x, y) = \begin{cases} C(x + 1 - \alpha, y) - C(1 - \alpha, y) & \text{if } \ 0 \le x \le \alpha \le 1 \\ y - C(1 - \alpha, y) + C(x - \alpha, y) & \text{if } \ 0 \le \alpha \le x \le 1. \end{cases}$$

Then copula U must satisfy the two functional equations

$$U(x + 1 - \alpha, y) = U(x, y) + U(1 - \alpha, y) \quad \text{if } 0 \le x \le \alpha \le 1 \text{ and} \tag{3.1}$$

$$U(x - \alpha, y) + y = U(x, y) + U(1 - \alpha, y) \quad \text{if } 0 \le \alpha \le x \le 1. \tag{3.2}$$

We will solve the above equations and show that the only copula which satisfies them is the product copula Π.

Fix $y \in [0,1]$ and let $f(x) = U(x,y)$. Then, from the properties of copulas, f is a continuous mapping on $[0,1]$ with boundary contitions $f(0) = 0$ and $f(1) = y$. Then (3.1) and (3.2) become

$$f(x+1-\alpha) = f(x) + f(1-\alpha) \quad \text{if } 0 \le x \le \alpha \le 1 \text{ and} \tag{3.3}$$

$$f(x-\alpha) + f(1) = f(x) + f(1-\alpha) \quad \text{if } 0 \le \alpha \le x \le 1. \tag{3.4}$$

First, we solve (3.3). Let $z = 1-\alpha$. Then (3.3) becomes the well-known Cauchy equation

$$f(x+z) = f(x) + f(z)$$

where $f\colon [0,1] \to [0,1]$ is continuous.

Observe that $f(\frac{m}{n}) = f(\frac{m-1}{n}) + f(\frac{1}{n})$ for all $m, n \in \mathbb{N}$ such that $1 \le m \le n$. Hence, by induction, we have $f(\frac{m}{n}) = mf(\frac{1}{n})$. Thus $f(1) = nf(\frac{1}{n})$. In other words, $f(\frac{1}{n}) = \frac{1}{n}f(1)$ for all $n \in \mathbb{N}$. Therefore $f(\frac{m}{n}) = \frac{m}{n}f(1)$ for all $m, n \in \mathbb{N}$ such that $1 \le m \le n$, i.e. $f(r) = rf(1)$ for all $r \in \mathbb{Q} \cap (0,1]$. We know that f is continuous and $f(0) = 0$. Hence, for all $x \in [0,1]$, we have

$$f(x) = xf(1). \tag{3.5}$$

Now, we solve (3.4). Observe that $f(x-x) + f(1) = f(x) + f(1-x)$. Hence, $f(1-x) = f(1) - f(x)$ for all $x \in [0,1]$. Thus, from (3.3), we have $f(x-\alpha) = f(x) - f(\alpha)$ for all $0 \le \alpha \le x \le 1$. In other words, $f(x) = f(x-\alpha) + f(\alpha)$ for all $0 \le \alpha \le x \le 1$. Again, we have $f(\frac{m}{n}) = f(\frac{m-1}{n}) + f(\frac{1}{n})$ for all $m, n \in \mathbb{N}$ such that $1 \le m \le n$. This is the same equation as the one we just solved. Hence, for all $x \in [0,1]$, we also have that

$$f(x) = xf(1). \tag{3.6}$$

From (3.5) and (3.6), we have $U(x, y) = f(x) = xf(1) = xy$ for all $x, y \in [0, 1]$. Thus, the only copula which satifies (3.1) and (3.2) is the product copula Π. $\square$

Lemma 3.15. *If* $\mathbf{C} \in \mathcal{M}$ *is a family such that* $*_{\mathbf{C}}$ *has a zero, then*

$$\int_0^1 C_t(x, y)\ dt = xy$$

for all $x, y \in [0, 1]$.

Proof. If $*_{\mathbf{C}}$ has the zero, then $\Pi(x, y) = (\Pi *_{\mathbf{C}} \Pi)(x, y) = \int_0^1 C_t(x, y)\ dt$. $\square$

Recall that $\{\Pi *_{\mathbf{C}} \Pi \colon \mathbf{C} \in \mathcal{M}\} = \mathfrak{C}$. Hence, $\Pi *_{\mathbf{C}} \Pi$ can be any copula. But for the $*_{\mathbf{C}}$ product to have a zero, $\Pi *_{\mathbf{C}} \Pi$ can only be the product copula Π. One can see that, for the product $*_{\mathbf{C}}$ to have a zero, the underlying family $\mathbf{C}$ must be extremely special.

Example 3.16. Given a copula $C \in \mathfrak{C}$. Let $\mathbf{C} = \{C\}_{t\in[0,1]}$. If $C = \Pi$, then the $*_{\mathbf{C}}$ product is simply the classical $*$-product, which has a zero. If $C \neq \Pi$, then the $*_{\mathbf{C}}$ product has no zero by the above Lemma.

Example 3.17. Let $\mathbf{C}$ be a family of copulas where $C_t = \Pi$ a.e. $t \in [0, 1]$. Then, $*_{\mathbf{C}}$ has a zero since the families $\mathbf{C}$ and $\{\Pi\}_{t\in[0,1]}$ are $*$-equivalent. In fact, $*_{\mathbf{C}}$ is identical to the classical $*$-product.

Example 3.18. Recall that the Farlie-Gumbel-Morgenstern copulas are of the form $C_\theta(u, v) = uv + \theta uv(1 - u)(1 - v)$ where $\theta \in [-1, 1]$. Let $\mathbf{C} = \{C_t\}_{t\in[0,1]}$ where C_t is equal to C_θ if $t \in [0, 1/2]$ and is equal to $C_{-\theta}$ otherwise. It is easily seen that the family $\mathbf{C}$ satisfies the condition in Lemma 3.15.

We will show that $*_{\mathbf{C}}$ in the above example has no zero, which implies that the criteria in Lemma 3.15 is not sufficient for the product to have a zero.

Example 3.19. Consider a family of copulas $\mathbf{C}$ in Example 3.18 where $\theta \neq 0$. Compute

$$\begin{aligned}
(A *_{\mathbf{C}} \Pi)(x,y) &= \int_0^1 C_t(\partial_2 A(x,t), y)\ dt \\
&= \int_0^{1/2} C_\theta(\partial_2 A(x,t), y)\ dt + \int_{1/2}^1 C_{-\theta}(\partial_2 A(x,t), y)\ dt \\
&= \int_0^{1/2} \partial_2 A(x,t)y + \theta\partial_2 A(x,t)y(1-\partial_2 A(x,t))(1-y)\ dt + \\
&\quad \int_{1/2}^1 \partial_2 A(x,t)y + \theta\partial_2 A(x,t)y(1-\partial_2 A(x,t))(1-y)\ dt \\
&= xy + \theta y(1-y)\Bigg[\int_0^{1/2} \partial_2 A(x,t)(1-\partial_2 A(x,t))\ dt - \\
&\quad \int_{1/2}^1 \partial_2 A(x,t)(1-\partial_2 A(x,t))\ dt.\Bigg]
\end{aligned}$$

Choose $A = C_\theta$. From straightforward computation, if $x \notin \{0,1\}$, then

$$\int_0^{1/2} \partial_2 A(x,t)(1-\partial_2 A(x,t))\ dt - \int_{1/2}^1 \partial_2 A(x,t)(1-\partial_2 A(x,t))\ dt = \frac{x^2}{2\theta x(x-1)} \neq 0.$$

Thus $C_\theta *_{\mathbf{C}} \Pi \neq \Pi$. Therefore, $*_{\mathbf{C}}$ has no zero.

We end this chapter with the list of open problems we encountered during our work on this part of thesis.

1. What are necessary and sufficient conditions for a family of copulas $\mathbf{C}$ to be in the set $\mathcal{M}$?

2. What are necessary and sufficient conditions for a family of copulas $\mathbf{C}$ to induce the product $*_{\mathbf{C}}$ which posesses a zero?

3. What are the invertible copulas with respect to the $*_{\mathbf{C}}$ product?

4. What are probabilistic interpretations of the $*_{\mathbf{C}}$ product?

REFERENCES

[1] W.F. Darsow, B. Nguyen, E.T. Olsen, Copulas and Markov processes, *Illinios J. Math.*, **36**(1992), 600-642.

[2] W.F. Darsow, E.T. Olsen, Norms for Copulas, *Int. J. Math. Math. Sci.*, **18(3)**(1995), 417-436.

[3] F. Durante, E.P. Klement, J.J. Quesada-Molina "Copulas: compatibility and Fréchet classes, "[Online] Available: `http://arxiv.org/abs/0711.2409v1`, (November 15, 2007).

[4] F. Durante, E.P. Klement, J.J. Quesada-Molina, P. Sarkoci, Remarks on two product-like constructions for copulas, *Kybernetika (Prague)*, **43(2)**(2007), 235-244.

[5] F. Durante, P. Sarkoci, C. Sempi, Shuffles of copulas, *J. Math. Anal. Appl.*, **352**(2009), 914-921.

[6] A.S. Kechris, Classical Descriptive Set Theory, Springer-Verlag, New York, 1995.

[7] R.B. Nelsen, An Introduction to Copulas, second ed., Springer Ser. Statist., Springer, New York, 2006.

[8] A. Rényi, On measures of dependence., *Acta. Math. Acad. Sci. Hungar.*, **10**(1959), 441-451.

[9] T. Santiwipanont, S. Sumetkijakan, Mutual complete dependence copulas and the $*$-product, preprint.

[10] B. Schweizer, E.F. Wolff, On nonparametric measures of dependence for random variables, *Ann. Statist.*, **9(4)**(1981), 879-885.

[11] K.F. Siburg, P.A. Stoimenov, A scalar product for copulas, *J. Math. Anal. Appl.*, **344**(2008), 429-439.

[12] K.F. Siburg, P.A. Stoimenov, A measure of mutual complete dependence, *Metrika*, **71**(2009), 239-251.

VITA

Name :Mr. Pongpol Ruankong

Date of Birth :18 December 1984

Place of Birth :Nakhonsawan, Thailand

Education :B.A. (Mathematics),University of Virginia, 2007