



**HAL**  
open science

# A class of smooth, possibly data-adaptive, nonparametric copula estimators and corresponding resampling techniques with applications

Bingqing Yi

► **To cite this version:**

Bingqing Yi. A class of smooth, possibly data-adaptive, nonparametric copula estimators and corresponding resampling techniques with applications. *Statistics [math.ST]*. Université de Pau et des Pays de l'Adour (UPPA), FRA., 2023. English. NNT : 2023PAUU3045 . tel-04631269

**HAL Id: tel-04631269**

**<https://theses.hal.science/tel-04631269v1>**

Submitted on 2 Jul 2024

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

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



Distributed under a Creative Commons Attribution 4.0 International License

**A class of smooth, possibly data-adaptive,  
nonparametric copula estimators and  
corresponding resampling techniques with  
applications**

by

**Bingqing Yi**

Thesis submitted in total fulfillment of the requirements of the joint degree of Doctor of Philosophy  
(PhD) of

**L'Université de Pau et des Pays de l'Adour**  
Laboratoire de Mathématiques et de leurs Applications

and

**The University of Melbourne**  
School of Mathematics and Statistics

September 2023

# Examining Committee

The following served on the Examining Committee for this thesis.

Jury president:

- Jean-David Fermanian: Professeur des universités, ENSAE Paris

Examiners and jury members:

- Jean-David Fermanian: Professeur des universités, ENSAE Paris
- Oliver Lopez: Professeur des universités, Sorbonne Université

Additional jury members:

- Ghislain Verdier : Maître de conférences, Université de Pau et des Pays de l'Adour
- Oliver Faugeras: Maître de conférences, Toulouse School of Economics

Supervisors :

- Ivan Kojadinovic : Professeur des universités, Université de Pau et des Pays de l'Adour
- Pavel Krupskiy: Assistant professor, The University of Melbourne
- Mark Holmes: Professor, The University of Melbourne

# Résumé

Les copules sont des outils mathématiques permettant de modéliser la dépendance entre les composantes d'un vecteur aléatoire et sont fréquemment utilisées dans des domaines tels que la finance, l'économie et la gestion du risque. Les chapitres 1 et 2 passent en revue les principaux résultats théoriques liées aux copules: y sont présentés leurs propriétés de base, les méthodes d'estimation les plus connues, le processus de copule empirique et des techniques de rééchantillonnage associées. Le chapitre 3 propose une classe d'estimateurs non paramétriques lissés de copules, potentiellement adaptatifs, qui contient les copules empiriques de Bernstein introduites par [Sancetta and Satchell \(2004\)](#) (et donc la copule empirique beta proposée par [Segers et al. \(2017\)](#)). En particulier, une sous-classe d'estimateurs qui s'avère uniformément plus précis que la copule empirique beta dans les expériences de Monte Carlo considérées a été identifiée. De plus, des conditions sous lesquelles les processus de copule empiriques associés convergent faiblement sont données. Le chapitre 4 propose deux techniques de rééchantillonnage pour la classe d'estimateurs considérée au chapitre 3. La première technique s'inspire des travaux de [Kiriliouk et al. \(2021\)](#) et peut être utilisée pour approcher les processus de copule empiriques associés dans le cas i.i.d. La seconde technique est une extension lisse du bootstrap à multiplicateurs dépendants proposé dans [Bücher and Kojadinovic \(2016\)](#) et peut être utilisée pour approcher les processus de copule empiriques associés dans le cas de données faiblement dépendantes. De plus, deux classes d'estimateurs lisses des dérivées partielles du premier ordre de la copule sont également étudiées théoriquement et empiriquement. Le dernier chapitre propose de nouvelles directions de recherche, telles que l'application des estimateurs étudiés et des techniques de rééchantillonnage associées à la détection de ruptures et à l'inférence pour des modèles factorielles de copules.

**Mots-clés:** statistique mathématique, copules, estimation adaptative, lissage, rééchantillonnage

# Abstract

Copulas are mathematical tools for modeling the dependence between the components of a random vector. They are frequently used in fields such as finance, economics, and risk management. Chapter 1 and 2 of this thesis provide a review of the main results in the study of copulas including their basic properties, estimation methods, the empirical copula process and appropriate resampling schemes for the latter. Chapter 3 proposes a broad class of smooth, possibly data-adaptive nonparametric copula estimators that contains empirical Bernstein copulas introduced by [Sancetta and Satchell \(2004\)](#) (and thus the empirical beta copula proposed by [Segers et al. \(2017\)](#)). A specific subclass that performs uniformly better than the empirical beta copula in Monte Carlo experiments is identified. Furthermore, conditions under which related sequential empirical copula processes converge weakly are provided. Chapter 4 proposes two resampling techniques for the class of estimators considered in Chapter 3. One technique builds up on the work of [Kiriliouk et al. \(2021\)](#) and can be used to bootstrap related empirical copula processes in the i.i.d. case. The other technique is a smooth extension of the dependent multiplier bootstrap proposed in [Bücher and Kojadinovic \(2016\)](#) and can be used to bootstrap related empirical copula processes in the sequential time series case. In addition, two classes of smooth estimators of the first-order partial derivatives of the copula are also theoretically and empirically studied. The last chapter discusses potential future research directions, such as applying the studied estimators and corresponding resampling techniques for change-point detection and inference in factor copula models.

**Keywords:** mathematical statistics, copula, data-adaptive estimation, smoothing, resampling

# Acknowledgements

First, I would like to offer my greatest gratitude to professor Ivan Kojadinovic for his excellent supervision during my entire PhD studies. His expertise and patience have been invaluable to me and have made me understand what it means to be a good researcher. I would also like to offer my thanks to professor Pavel Krupskiy for the many discussions we had on factor copula models, and to Mark Holmes for the fruitful discussions on my first project with Ivan.

Furthermore, I am grateful to my defense committee for offering their valuable knowledge and expertise. Additionally, the generous support of Centre National de la Recherche Scientifique (CNRS) in funding my research was instrumental in making my work achievable.

Finally, I must express my very profound gratitude to my wife, Zhuoqing Gu, for providing me with unconditional support and continuous encouragement throughout my years of study and through the process of researching and writing this thesis. This accomplishment would not have been possible without her.

# Contents

<b>Examining Committee</b>	<b>i</b>
<b>Résumé</b>	<b>ii</b>
<b>Abstract</b>	<b>iii</b>
<b>Acknowledgements</b>	<b>ix</b>
<b>1 Introduction</b>	<b>1</b>
<b>2 A brief review on copulas and empirical copula processes</b>	<b>5</b>
2.1 Copulas . . . . .	6
2.1.1 Definition and basic properties . . . . .	6
2.1.2 Survival copulas . . . . .	9
2.1.3 Copula densities . . . . .	10
2.1.4 Families of copulas . . . . .	10
2.1.4.1 Archimedean copulas . . . . .	10
2.1.4.2 Elliptical copulas . . . . .	11
2.1.4.3 Extreme-value copulas . . . . .	12
2.2 Estimators of copulas . . . . .	12
2.2.1 Parametric estimators . . . . .	13
2.2.2 Semiparametric estimators . . . . .	14
2.2.3 Nonparametric estimators . . . . .	15
2.2.3.1 The empirical copula . . . . .	16
2.2.3.2 The empirical checkerboard copula . . . . .	17
2.2.3.3 The empirical Bernstein copula and the empirical beta copula . . . . .	18
2.3 Empirical copula processes . . . . .	18
2.3.1 Mathematical preliminaries . . . . .	19
2.3.2 Empirical copula processes . . . . .	21
2.3.2.1 Asymptotics in the i.i.d. case . . . . .	22
2.3.2.2 Asymptotics in the time series case . . . . .	23
2.3.3 The empirical Bernstein copula process and the empirical beta copula process . . . . .	26
2.4 Bootstrapping empirical copula processes . . . . .	27
2.4.1 Mathematical preliminaries . . . . .	27
2.4.2 Bootstrapping the empirical copula processes . . . . .	28
2.4.2.1 Straightforward bootstraps in the i.i.d. case . . . . .	28
2.4.2.2 Smooth bootstraps in the i.i.d. case . . . . .	29

2.4.2.3	Sequential dependent multiplier bootstrap for $\mathbb{C}_n$ in the time series case	30
<b>3</b>	<b>A class of smooth, possibly data-adaptive nonparametric copula estimators containing the empirical beta copula</b>	<b>34</b>
3.1	Introduction	35
3.2	A broad class of smooth, possibly data-adaptive empirical copulas	37
3.3	Properties and general form of the smooth estimators	40
3.3.1	Univariate margins of the smooth estimators	40
3.3.2	General form of the smooth estimators	41
3.3.3	Conditions for being multivariate d.f.s	43
3.4	Estimators based on smoothing distributions with scaled binomial, scaled beta-binomial or beta margins	43
3.4.1	Scaled binomial and beta-binomial smoothing survival margins	44
3.4.2	Beta smoothing survival margins	46
3.4.3	On the survival copula of the smoothing distributions	48
3.5	Suggested data-adaptive alternatives to the empirical beta copula	49
3.5.1	Experiments for understanding the influence of the spread and the shape of the smoothing distributions	49
3.5.2	Suggested data-adaptive smooth estimators and their finite-sample performance	52
3.6	Asymptotics of related sequential empirical processes	57
3.7	Concluding remarks	60
3.8	Proofs	61
3.8.1	Proofs of Lemma 3.2.1, Proposition 3.3.2, Proposition 3.3.7 and Proposition 3.3.10	61
3.8.2	Proofs of Propositions 3.4.1, 3.4.2 and 3.4.4	63
3.8.3	Proofs of Theorem 3.6.4	66
<b>4</b>	<b>Resampling techniques for a class of smooth, possibly data-adaptive empirical copulas</b>	<b>75</b>
4.1	Introduction	76
4.2	Smooth, possibly data-adaptive, empirical copulas and their asymptotics	79
4.2.1	Smooth, possibly data-adaptive, nonparametric copula estimators	79
4.2.2	Asymptotics of related sequential processes	82
4.3	Bootstrap by drawing samples from the estimators in the i.i.d. case	84
4.3.1	Drawing samples from the smooth empirical copulas	85
4.3.2	Asymptotic validity results	87
4.3.3	Finite-sample comparison of two smooth bootstraps	88
4.4	Smooth sequential dependent multiplier bootstraps in the time series case	90
4.4.1	Main intuition and existing work	90
4.4.2	I.i.d. and dependent multiplier sequences	91
4.4.3	Non-smooth sequential dependent multiplier replicates	91
4.4.4	Smooth sequential dependent multiplier replicates	92
4.4.5	Finite-sample comparison of three multiplier bootstraps	94
4.4.6	Application to change-point detection	99
4.5	Estimators of the first-order partial derivatives of the copula	102
4.5.1	Estimators based on finite differences of the empirical copula	103



---

4.5.2	Two classes of smooth estimators . . . . .	104
4.5.3	Weak consistency . . . . .	106
4.5.4	Finite-sample performance of selected estimators . . . . .	108
4.6	Conclusion . . . . .	112
4.7	Proofs . . . . .	113
4.7.1	Proof of Corollary 4.2.13 . . . . .	113
4.7.2	Proofs of Proposition 4.3.3 and Lemma 4.7.4 . . . . .	114
4.7.3	Proof of Theorem 4.3.5 . . . . .	116
4.7.4	Proof of Theorem 4.4.3 . . . . .	120
4.7.5	Proofs of Propositions 4.5.3, 4.5.5, 4.5.9 and 4.5.10 . . . . .	122
<b>5</b>	<b>Conclusion and some future research directions</b>	<b>133</b>
5.1	Conclusion . . . . .	133
5.2	Some future research directions . . . . .	134
5.2.1	Application to sequential change-point detection . . . . .	134
5.2.2	Application to high-dimensional factor copula models . . . . .	135

# Chapter 1

## Introduction

The study of how random variables are related to each other is a crucial aspect of probability theory and statistics. Over time, various ideas and concepts have been developed to describe these relationships. The most widely used approach is the *Pearson correlation coefficient*, also known as *Pearson's r*. However, it is widely recognized that this method only captures linear relationships, and is only able to fully describe a limited number of multivariate distribution types, such as the multivariate normal distribution. Rank-based correlation were introduced next. For instance, [Spearman \(1904\)](#) proposed *Spearman's rho* and [Kendall \(1938\)](#) proposed *Kendall's tau*. However, the aforementioned approaches are not able to describe in full the dependence between two random variables since they only provide information about the overall degree of dependence. [Sklar \(1959\)](#) proposed a theorem that highlights the existence of mathematical objects known as copulas that fully describe the dependence between the components of a random vector.

In short, a copula  $C$  of a random vector  $\mathbf{X} = (X_1, \dots, X_d)$  is a function from  $[0, 1]^d$  to  $[0, 1]$  that captures the relationship between the components of  $\mathbf{X}$ . Specifically, let  $F$  be the multivariate distribution function of  $\mathbf{X}$  and, for  $j \in \{1, \dots, d\}$ , let  $F_j$  be the univariate distribution function of  $X_j$ . Then, Sklar's theorem states that there exists a copula  $C$  such that

$$F(\mathbf{x}) = C\{F_1(x_1), \dots, F_d(x_d)\}, \quad \mathbf{x} = (x_1, \dots, x_d) \in \mathbb{R}^d. \quad (1.1)$$

Moreover, when  $X_1, \dots, X_d$  are continuous univariate random variables,  $C$  is unique. The central implication of this identity, with mathematical details given in later chapters, can be heuristically summarized as follows: a copula is a function that captures the whole and the only information about the stochastic dependence between the components of  $\mathbf{X}$ . As a consequence of (1.1),  $F$  can be modeled in two separate steps: the first step consists of estimating the univariate margins  $F_1, \dots, F_d$  and the second step consists of estimating the copula  $C$ . This work is only concerned with the estimation of the copula.

For the practitioner's convenience, various parametric families have been proposed. The most commonly used ones are, among others, *Archimedean copulas*, *elliptical copulas* and *extreme-value copulas* (see, e.g., Joe, 1997; Nelsen, 2006; Durante and Sempi, 2010). However, before carrying out statistical inference on the unknown copula  $C$ , one often needs to choose the appropriate parametric family. For this purpose, it is useful to first have at hand a nonparametric estimator of  $C$ . The best known one is the *empirical copula* (Rüschendorf, 1976; Deheuvels, 1979), which has been used widely in inference procedures on the unknown copula  $C$ . For instance, it is used to test the symmetry and exchangeability of a copula before choosing the appropriate family (see, e.g. Kojadinovic and Yan, 2010; Genest et al., 2012; Krupskii, 2017). The empirical copula has however an obvious drawback: it is a step function, while a genuine copula is a Lipschitz continuous function (see Proposition 2.1.6 in Section 2.1.1). As a result, its estimation bias is typically large when the sample size is small. To overcome this drawback, various smooth estimators have been proposed, such as the Kernel-based copula estimators (Fermanian and Scaillet, 2003; Gribkova and Lopez, 2015), the *empirical checkerboard copula* (Genest et al., 2017; Genest and Nešlehová, 2007) and the *empirical beta copula* (Segers et al., 2017) (which is a special case of the *empirical Bernstein copula* (Sancetta and Satchell, 2004; Janssen et al., 2012)). The numerical study in Segers et al. (2017) shows that the empirical beta copula outperforms the empirical copula, the empirical checkerboard copula and, in several cases, the empirical Bernstein copula with the smoothing rate suggested by Janssen et al. (2012).

For the asymptotic study of many existing inference procedures on the unknown copula  $C$ , a key ingredient is the so-called *empirical copula process*, which has been studied by Fermanian et al. (2004), van der Vaart and Wellner (2007), Segers (2012) and Bücher and Volgushev (2013) among others. Furthermore, a more general *two sided empirical copula process* was proposed and studied in Bücher and Kojadinovic (2016). The (sequential) empirical copula process is often used to study the asymptotics of the test statistics based on the empirical copula, for instance, for the test of independence in Genest and Rémillard (2004) and the *change-point detection* test in Bücher et al. (2014). In addition, note that the *empirical Bernstein copula process*, which is a smooth version of the empirical copula process, was proposed and studied in Segers et al. (2017).

In terms of asymptotics, the (smooth sequential) empirical copula process converges weakly to a Gaussian process with covariance structure depending on the the unknown copula, its first-order derivatives and, in the time series setting, the serial dependence properties of the data sequence (see, e.g. Fermanian et al., 2004; Segers, 2012; Bücher and Volgushev, 2013; Bücher and Kojadinovic, 2016; Segers et al., 2017). Therefore in almost all inference procedures, it is necessary to rely on resampling techniques to compute corresponding confidence intervals or p-values. Apart from the most classic (multinomial) bootstrap, one frequently used resampling technique in the literature is the so-called *multiplier bootstrap* (see, e.g., Scaillet, 2005; Rémillard and Scaillet, 2009). For the purpose of detecting cross-sectional changes in a multivariate time series, the multiplier bootstrap was extended to the time series and sequential settings in Bücher and Kojadinovic (2016) and Bücher et al. (2014). Another resampling method is

the so-called *smooth bootstrap* based on drawing samples from the empirical beta copula. It was recently studied both theoretically and empirically in [Kiriliouk et al. \(2021\)](#). The Monte Carlo experiments reported therein reveal that it is a competitive alternative to the multiplier bootstrap while being substantially simpler to implement. However, one practical inconvenience is that the aforementioned smooth bootstrap cannot be directly extended to the time series setting.

This work is organized as follows: Chapter 2 offers a review on some of the major results in the study of copulas. We state the definition of a copula with some of its basic properties in a more precise and rigorous way, and provide a brief review of various estimators including parametric estimators, semiparametric estimators and nonparametric estimators. Additionally, we also make a summary of the study of various *(smooth) empirical copula processes* and their corresponding bootstraps. From a historical perspective, the theoretical developments for the latter can be summarized as follows: from i.i.d. to serially dependent, from nonsequential to sequential and from nonsmooth to smooth.

Chapter 3 and 4 include the main contributions. In Chapter 3, we propose a broad class of smooth, possibly data-adaptive nonparametric copula estimators, which are extensions of the smooth estimators in [Sancetta and Satchell \(2004\)](#) and [Segers et al. \(2017\)](#). We provide conditions under which the estimate is a genuine copula and, by extending the results in [Segers et al. \(2017\)](#) and [Bücher and Kojadinovic \(2016\)](#), we provide conditions under which the related smooth sequential empirical copula process converges weakly. In Chapter 4, we propose two resampling schemes to obtain the replicates of the latter process and study their asymptotic validity. The first scheme, following [Kiriliouk et al. \(2021\)](#), involves drawing samples from the smooth estimator and can be used in the nonsequential i.i.d. case. In particular, the algorithm to sample from it can be viewed as an extension of Algorithm 3.2 in [Kiriliouk et al. \(2021\)](#). The second scheme is a smooth extension of the sequential dependent multiplier bootstrap in [Bücher and Kojadinovic \(2016\)](#) and can be used in the sequential time series case. As far as the Monte Carlo experiments are concerned, it is shown in Chapter 3 that one particular subclass of estimators is found to be uniformly better than the empirical beta copula in [Segers et al. \(2017\)](#), and in Chapter 4 that the two resampling schemes can offer advantages over the corresponding schemes in [Kiriliouk et al. \(2021\)](#) and [Bücher and Kojadinovic \(2016\)](#). Additionally, Chapter 4 proposes two classes of smooth estimators of the first-order partial derivatives of a copula, with some estimators displaying substantially better finite-sample performance than the Bernstein estimator studied in [Janssen et al. \(2016\)](#). Chapter 3 corresponds to [Kojadinovic and Yi \(2022\)](#) accepted to the Journal of Multivariate Analysis and Chapter 4 corresponds to [Kojadinovic and Yi \(2023\)](#) currently under submission.

Finally, after summarizing our contributions, we provide, in Chapter 5, two directions of extension that could lead to future projects. The first one can be built up on [Kojadinovic and Verdier \(2021\)](#) and involves applying the smooth empirical copula processes in Chapter 3 and the corresponding multiplier bootstraps in Chapter 4 to *sequential change-point detection*. The goal is to detect changes in the cross-sectional dependence of multivariate time series when new

---

observations are obtained sequentially. The second one can be built up on [Kurpskii and Joe \(2020\)](#) and involves applying the estimators of the first-order partial derivatives of a copula introduced in Chapter 4 to the *one-factor copula model* proposed in [Kurpskii and Joe \(2013\)](#) and studied in the high-dimensional case in [Kurpskii and Joe \(2020\)](#).

# Chapter 2

## A brief review on copulas and empirical copula processes

2.1	Copulas . . . . .	6
2.1.1	Definition and basic properties . . . . .	6
2.1.2	Survival copulas . . . . .	9
2.1.3	Copula densities . . . . .	10
2.1.4	Families of copulas . . . . .	10
2.1.4.1	Archimedean copulas . . . . .	10
2.1.4.2	Elliptical copulas . . . . .	11
2.1.4.3	Extreme-value copulas . . . . .	12
2.2	Estimators of copulas . . . . .	12
2.2.1	Parametric estimators . . . . .	13
2.2.2	Semiparametric estimators . . . . .	14
2.2.3	Nonparametric estimators . . . . .	15
2.2.3.1	The empirical copula . . . . .	16
2.2.3.2	The empirical checkerboard copula . . . . .	17
2.2.3.3	The empirical Bernstein copula and the empirical beta copula . . . . .	18
2.3	Empirical copula processes . . . . .	18
2.3.1	Mathematical preliminaries . . . . .	19

2.3.2	Empirical copula processes . . . . .	21
2.3.2.1	Asymptotics in the i.i.d. case . . . . .	22
2.3.2.2	Asymptotics in the time series case . . . . .	23
2.3.3	The empirical Bernstein copula process and the empirical beta copula process . . . . .	26
2.4	Bootstrapping empirical copula processes . . . . .	27
2.4.1	Mathematical preliminaries . . . . .	27
2.4.2	Bootstrapping the empirical copula processes . . . . .	28
2.4.2.1	Straightforward bootstraps in the i.i.d. case . . . . .	28
2.4.2.2	Smooth bootstraps in the i.i.d. case . . . . .	29
2.4.2.3	Sequential dependent multiplier bootstrap for $\mathbb{C}_n$ in the time series case . . . . .	30

This chapter aims to provide the background needed for the reader to understand our main contributions in Chapter 3 and 4. The notation, definitions and theorems in this sections mainly follow from Hofert et al. (2018), Durante and Sempi (2010), Durante and Sempi (2015), van der Vaart and Wellner (2000) and Kosorok (2008).

## 2.1 Copulas

In this section, we introduce copulas, survival copulas, copula densities, and we present some of their basic properties. Furthermore, we provide some examples of families of copulas.

### 2.1.1 Definition and basic properties

First, let us recall that the multivariate distribution function (d.f.) of a  $d$ -dimensional random vector  $\mathbf{X} = (X_1, \dots, X_d)$  is the function defined by

$$F(\mathbf{x}) = \mathbb{P}(\mathbf{X} \leq \mathbf{x}) = \mathbb{P}(X_1 \leq x_1, \dots, X_d \leq x_d), \quad \mathbf{x} = (x_1, \dots, x_d) \in \mathbb{R}^d. \quad (2.1)$$

Then, the univariate d.f.  $F_j$  of  $X_j$ ,  $j \in \{1, \dots, d\}$ , can be recovered from the multivariate d.f.  $F$  by  $F_j(x_j) = F(\infty, \dots, \infty, x_j, \infty, \dots, \infty)$ ,  $x_j \in \mathbb{R}$ . Copulas are particular multivariate d.f.s given by the definition below.

**Definition 2.1.1** (Copula). A *copula* is a multivariate d.f. with standard uniform univariate margins.

In the rest of this document, for any univariate d.f.  $H$ , let  $\text{ran } H$  denote the *range* of  $H$  defined by  $\text{ran } H = \{H(x) : x \in \mathbb{R}\}$  and let  $H^{-1}$  denote its associated quantile function (generalized inverse) defined by  $H^{-1}(y) = \inf\{x \in \mathbb{R} : H(x) \geq y\}$ ,  $y \in [0, 1]$ , with the convention that  $\inf \emptyset = \infty$ .

**Theorem 2.1.2** (Sklar (1959)). *For any  $d$ -dimensional d.f.  $F$  with univariate margins  $F_1, \dots, F_d$ , there exists a  $d$ -dimensional copula  $C$  such that,*

$$F(\mathbf{x}) = C \{F_1(x_1), \dots, F_d(x_d)\}, \quad \mathbf{x} \in \mathbb{R}^d. \quad (2.2)$$

The copula  $C$  is uniquely defined on  $\prod_{j=1}^d \text{ran } F_j$  and given by

$$C(\mathbf{u}) = F \{F_1^{-1}(u_1), \dots, F_d^{-1}(u_d)\}, \quad \mathbf{u} \in \prod_{j=1}^d \text{ran } F_j, \quad (2.3)$$

Conversely, given a copula  $C$  and univariate d.f.s  $F_1, \dots, F_d$ ,  $F$  defined by (2.2) is a multivariate d.f. with margins  $F_1, \dots, F_d$ .

**Lemma 2.1.3** (Probability integral transformation). *Let  $X$  be a univariate random variable with continuous d.f.  $H$ . Then  $H(X)$  is a standard uniform random variable.*

**Lemma 2.1.4** (Stochastic Analog of Sklar's Theorem). *Let  $\mathbf{X}$  be a  $d$ -dimensional random vector with continuous univariate marginal d.f.s  $F_1, \dots, F_d$ . Then  $\mathbf{X}$  has copula  $C$  if and only if the random vector  $\mathbf{U} = (F_1(X_1), \dots, F_d(X_d))$  has d.f.  $C$ .*

Then, the intuition of modeling dependence by copulas can be briefly summarized as follows: suppose we have two  $d$ -dimensional random vectors  $\mathbf{X} = (X_1, \dots, X_d)$  (resp.  $\mathbf{Y} = (Y_1, \dots, Y_d)$ ) with d.f.  $F$  (resp.  $G$ ) and marginal d.f.s  $F_1, \dots, F_d$  (resp.  $G_1, \dots, G_d$ ). For the moment, we assume that  $F_1, \dots, F_d, G_1, \dots, G_d$  are continuous. We wish to know whether the dependence structure between the components of  $\mathbf{X}$  is the same as that of  $\mathbf{Y}$ . We cannot compare  $F$  and  $G$  directly since they may have different marginal d.f.s and are therefore not “comparable”; see Chapter 1 in Hofert et al. (2018) for a motivating example showing how marginal d.f.s could affect one's perception on the measure of dependence. However, with Theorem 2.1.2 and Lemma 2.1.4 in mind, we can equivalently compare whether the random vector  $(F_1(X_1), \dots, F_d(X_d))$  and  $(G_1(Y_1), \dots, G_d(Y_d))$  have the same distribution or, in other words, whether  $\mathbf{X}$  and  $\mathbf{Y}$  have the same copula.

In the rest of this document, for any  $j \in \{1, \dots, d\}$  and any  $\mathbf{u} \in [0, 1]^d$ , let  $\mathbf{u}^{(j)}$  be the vector of  $[0, 1]^d$  defined by  $u_i^{(j)} = u_j$  if  $i = j$  and 1 otherwise. The proof of the following theorem can be found in Durante and Sempi (2015).

**Theorem 2.1.5.** *A function  $C : [0, 1]^d \mapsto [0, 1]$  is a copula if and only if the following conditions hold:*

1. (grounded)  $C(u_1, \dots, u_d) = 0$  if  $u_j = 0$  for at least one index  $j \in \{1, \dots, d\}$ .



2. (uniform margin)  $C(\mathbf{u}^{(j)}) = u_j$  for all  $\mathbf{u} \in [0, 1]^d$  and  $j \in \{1, \dots, d\}$ .
3. ( $d$ -increasing) For any  $\mathbf{a}, \mathbf{b} \in [0, 1]^d$  and  $\mathbf{a} \leq \mathbf{b}$ ,

$$\sum_{\mathbf{i} \in \{0,1\}^d} (-1)^{\sum_{j=1}^d i_j} C(a_1^{i_1} b_1^{1-i_1}, \dots, a_d^{i_d} b_d^{1-i_d}) \geq 0,$$

where the summation is taken over all  $2^d$ -many vectors  $\mathbf{i} = (i_1, \dots, i_d)$  for  $i_1, \dots, i_d \in \{0, 1\}$ .

Next, we present some basic properties of copulas, the proof of which can be found in Joe (1997) and Nelsen (2006).

**Proposition 2.1.6** (Basic properties of copulas). *Let  $C$  be a  $d$ -dimensional copula and suppose that  $\mathbf{X} = (X_1, \dots, X_d)$  is a random vector with marginal d.f.s  $F_1, \dots, F_d$  and copula  $C$ . Then the following results hold.*

1. (Fréchet-Hoeffding bounds) For any copula  $C$ , we have that

$$W(\mathbf{u}) \leq C(\mathbf{u}) \leq M(\mathbf{u}), \quad \mathbf{u} \in [0, 1]^d,$$

where

$$W(\mathbf{u}) = \max \left\{ \sum_{i=1}^d u_i - d + 1, 0 \right\}, \quad M(\mathbf{u}) = \min\{u_1, \dots, u_d\}, \quad \mathbf{u} \in [0, 1]^d.$$

It is important to note that  $W$  is a copula only if  $d = 2$  whereas  $M$  is a copula for all  $d \geq 2$ .

2. (Independent copula) If  $F_1, \dots, F_d$  are continuous, then  $X_1, \dots, X_d$  are mutually independent if and only if

$$C(\mathbf{u}) = \prod_{j=1}^d u_j, \quad \mathbf{u} \in [0, 1]^d.$$

3. (Lipschitz-continuity)  $C$  is Lipschitz continuous with respect to the  $L^1$  norm on  $[0, 1]^d$ . Namely,

$$|C(\mathbf{u}) - C(\mathbf{v})| \leq \sum_{j=1}^d |u_j - v_j|, \quad \text{for all } \mathbf{u}, \mathbf{v} \in [0, 1]^d.$$

4. (Invariance under increasing transformations) If  $F_1, \dots, F_d$  are continuous and  $H_1, \dots, H_d$  are strictly increasing, then  $(H_1(X_1), \dots, H_d(X_d))$  also has copula  $C$ .

5. (Kendall's tau and Spearman's rho) If  $d = 2$  and  $F_1$  and  $F_2$  are continuous, then Kendall's tau and Spearman's rho of  $X_1$  and  $X_2$  can be expressed as

$$\begin{aligned}\tau(X_1, X_2) &= 4 \int_{[0,1]^2} C(\mathbf{u}) dC(\mathbf{u}) - 1, \\ \rho(X_1, X_2) &= 12 \int_{[0,1]^2} C(\mathbf{u}) d\mathbf{u} - 3.\end{aligned}$$

### 2.1.2 Survival copulas

Let us recall that the *multivariate survival function*  $\bar{F}$  corresponding to the multivariate d.f.  $F$  in (2.1) is the function defined by

$$\bar{F}(\mathbf{x}) = \mathbb{P}(\mathbf{X} > \mathbf{x}) = \mathbb{P}(X_1 > x_1, \dots, X_d > x_d), \quad \mathbf{x} = (x_1, \dots, x_d) \in \mathbb{R}^d. \quad (2.4)$$

Then, the *univariate survival function*  $\bar{F}_j$  of  $X_j$ ,  $j \in \{1, \dots, d\}$ , can be recovered from the multivariate survival function  $\bar{F}$  by  $\bar{F}_j(x_j) = \bar{F}(-\infty, \dots, -\infty, x_j, -\infty, \dots, -\infty)$ ,  $x_j \in \mathbb{R}$ .

**Theorem 2.1.7** (Sklar's Theorem for survival functions). *For any  $d$ -dimensional survival function  $\bar{F}$  with margins  $\bar{F}_1, \dots, \bar{F}_d$ , there exists a copula  $\bar{C}$  such that*

$$\bar{F}(\mathbf{x}) = \bar{C} \{ \bar{F}_1(x_1), \dots, \bar{F}_d(x_d) \}, \quad \mathbf{x} \in \mathbb{R}^d, \quad (2.5)$$

The copula  $\bar{C}$  is uniquely defined on  $\prod_{j=1}^d \text{ran } \bar{F}_j$  and given by

$$\bar{C}(\mathbf{u}) = \bar{F} \{ \bar{F}_1^{-1}(u_1), \dots, \bar{F}_d^{-1}(u_d) \}, \quad \mathbf{u} \in \prod_{j=1}^d \text{ran } \bar{F}_j. \quad (2.6)$$

Conversely, given any  $d$ -dimensional copula  $\bar{C}$  and univariate survival functions  $\bar{F}_1, \dots, \bar{F}_d$ ,  $\bar{F}$  defined in (2.5) is a survival function with margins  $\bar{F}_1, \dots, \bar{F}_d$ .

**Proposition 2.1.8** (Properties of survival copulas). *Let  $C$  be a  $d$ -dimensional copula and let  $\mathbf{U}$  be a random vector whose d.f. is  $C$ . Then,*

1.  $(1 - U_1, \dots, 1 - U_d)$  is a random vector whose d.f. is the survival copula  $\bar{C}$  of  $C$ .
2. The survival copula  $\bar{C}$  can be computed from  $C$  via

$$\bar{C}(\mathbf{u}) = \sum_{J \subset \{1, \dots, d\}} (-1)^{|J|} C \left( (1 - u_1)^{\mathbf{1}(1 \in J)}, \dots, (1 - u_d)^{\mathbf{1}(d \in J)} \right), \quad \mathbf{u} \in [0, 1]^d, \quad (2.7)$$

where the sum extends over all  $2^d$  subsets  $J$  of  $\{1, \dots, d\}$ ,  $|J|$  denotes the number of elements of  $J$  and  $\mathbf{1}(j \in J)$  is the indicator of  $j \in J$ .

### 2.1.3 Copula densities

A copula  $C$  is called *absolutely continuous* if it admits a density. In this work, we follow Section 2.1 in Hofert et al. (2018) and shall simply say that a copula  $C$  admits a *density*  $c$  if

$$c(\mathbf{u}) = \frac{\partial^d}{\partial u_d \dots \partial u_1} C(u_1, \dots, u_d), \quad \mathbf{u} \in (0, 1)^d, \quad (2.8)$$

exists and is integrable. In that case, the density  $f$  of  $F$  can be expressed as

$$f(\mathbf{x}) = c\{F_1(x_1), \dots, F_d(x_d)\} \prod_{j=1}^d f_j(x_j), \quad \mathbf{x} \in \prod_{j=1}^d \text{ran } X_j, \quad (2.9)$$

where, for any  $j \in \{1, \dots, d\}$ ,  $\text{ran } X_j$  is the range of  $X_j$ ,  $f_j$  denotes the density of  $F_j$  and  $c$  denotes the density of  $C$ . Note that if  $C$  admits a density  $c$ , so does  $\bar{C}$  and the density  $\bar{c}$  of  $\bar{C}$  is given by

$$\bar{c}(\mathbf{u}) = c(1 - u_1, \dots, 1 - u_d), \quad \mathbf{u} \in (0, 1)^d.$$

### 2.1.4 Families of copulas

Analogous to various families of d.f.s, there are various parametric families of copulas that are commonly used in statistical modeling, each with its own unique properties and characteristics. Here we introduce the three most commonly used ones.

#### 2.1.4.1 Archimedean copulas

**Definition 2.1.9** (Archimedean copulas). A  $d$ -dimensional copula  $C$  is called *Archimedean* if it admits the representation

$$C(\mathbf{u}) = \psi(\psi^{-1}(u_1) + \dots + \psi^{-1}(u_d)), \quad \mathbf{u} \in [0, 1]^d,$$

where  $\psi : [0, \infty] \mapsto [0, 1]$  is the so-called *Archimedean generator* that satisfies the following conditions

1.  $\psi(0) = 1$  and  $\psi(\infty) = \lim_{t \rightarrow \infty} \psi(t) = 0$ .
2.  $\psi$  is continuous and decreasing on  $[0, \infty]$  and strictly decreasing on  $[0, \inf\{t : \psi(t) = 0\}]$ .
3.  $\psi$  admits derivatives  $\psi^{(k)}$  up to the order  $k = d - 2$  satisfying  $(-1)^k \psi^{(k)}(t) \geq 0$  for all  $k \in \{0, \dots, d - 2\}$ ,  $t \in (0, \infty)$ , and  $t \mapsto (-1)^{d-2} \psi^{(d-2)}(t)$  is decreasing and convex on  $(0, \infty)$ .

**Example 2.1.10** (Examples of bivariate Archimedean copulas). By specifying  $\psi$ , we are able to generate different classes of copulas.

- The independence copula: Let  $\psi(t) = -\log(t)$ , then

$$C(u, v) = uv, \quad (u, v) \in [0, 1]^2.$$

- Ali–Mikhail–Haq copulas [Ali et al. (1978)]: for  $\theta \in [-1, 1]$ , let  $\psi_\theta(t) = \log \left\{ \frac{1-\theta(1-t)}{t} \right\}$ , then

$$C_\theta(u, v) = \frac{uv}{1 - \theta(1-u)(1-v)}, \quad (u, v) \in [0, 1]^2.$$

- Clayton copulas [Clayton (1978)]: for  $\theta \in [-1, 0) \cup (0, \infty)$ , let  $\psi_\theta(t) = \frac{1}{\theta}(t^{-\theta} - 1)$ , then

$$C_\theta(u, v) = \left\{ \max(u^{-\theta} + v^{-\theta} - 1, 0) \right\}^{-1/\theta}, \quad (u, v) \in [0, 1]^2.$$

- Frank copulas [Genest (1987)]: for  $\theta \in (-\infty, 0) \cup (0, \infty)$ , let  $\psi_\theta(t) = -\log \left( \frac{\exp(-\theta t) - 1}{\exp(-\theta) - 1} \right)$ , then

$$C_\theta(u, v) = -\frac{1}{\theta} \log \left[ 1 + \frac{\{\exp(-\theta u) - 1\} \{\exp(-\theta v) - 1\}}{\{\exp(-\theta) - 1\}} \right], \quad (u, v) \in [0, 1]^2.$$

- Gumbel–Hougaard copulas [Gumbel (1961), Hougaard (1986)]: for  $\theta \in [1, \infty)$ , let  $\psi_\theta(t) = \{-\log(t)\}^\theta$ , then

$$C_\theta(u, v) = \exp \left[ - \left\{ (-\log(u))^\theta + (-\log(v))^\theta \right\}^{-\frac{1}{\theta}} \right], \quad (u, v) \in [0, 1]^2.$$

- Joe copulas [Joe (1997)]: for  $\theta \in [1, \infty)$ , let  $\psi_\theta(t) = -\log \{1 - (1-t)^\theta\}$ , then

$$C_\theta(u, v) = 1 - \left\{ (1-u)^\theta + (1-v)^\theta - (1-u)^\theta(1-v)^\theta \right\}^{1/\theta}, \quad (u, v) \in [0, 1]^2.$$

#### 2.1.4.2 Elliptical copulas

**Definition 2.1.11** (Elliptical distributions). A  $d$ -dimensional random vector  $\mathbf{X}$  is said to have an *elliptical distribution* with mean  $\mathbf{u}$ , covariance matrix  $\Sigma = (\sigma_{ij})$  and generator  $g : [0, \infty) \rightarrow [0, \infty)$ , and one writes  $\mathbf{X} \sim \mathcal{E}(\mathbf{u}, \Sigma, g)$ , if it can be expressed in the form

$$\mathbf{X} = \mathbf{u} + R\mathbf{A}\mathbf{U},$$

where  $\mathbf{A}\mathbf{A}^\top = \Sigma$  is the Cholesky decomposition of  $\Sigma$ ,  $\mathbf{U}$  is a  $d$ -dimensional random vector uniformly distributed on the sphere  $\{\mathbf{u} \in \mathbb{R}^d : \sum_{j=1}^d u_j^2 = 1\}$  and  $R$  is a positive random variable independent of  $\mathbf{U}$ , with density given, for every  $r > 0$ , by

$$f_g(r) = \frac{2\pi^{d/2}}{\Gamma(d/2)} r^{d-1} g(r^2). \quad (2.10)$$

**Example 2.1.12** (Examples of elliptical distributions). Here we provide two examples of elliptical distribution.

- If  $g(t) = (2\pi)^{-d/2} \exp(-t/2)$ , then  $\mathbf{X}$  has a multivariate Gaussian distribution.
- If  $g(t) = c(1 + t/v)^{-(d+v)/2}$  for a suitable constant  $c$ , then  $\mathbf{X}$  has a multivariate t-Student distribution with  $v$  degrees of freedom.

**Definition 2.1.13** (Elliptical copulas). Let  $\mathbf{X}$  be a  $d$ -dimensional elliptical random vector,  $\mathbf{X} \sim \mathcal{E}(\mathbf{u}, \Sigma, g)$ . Suppose that, for every  $j \in \{1, \dots, d\}$ ,  $(X_j / \sqrt{\sigma_{jj}})$  has d.f.  $F_g$ , where  $F_g$  is the d.f. of  $f_g$  in (2.10) when  $d = 1$ . Then, we call *elliptical copula* the d.f. of the random vector

$$\left( F_g \left( \frac{X_1}{\sqrt{\sigma_{11}}} \right), \dots, F_g \left( \frac{X_d}{\sqrt{\sigma_{dd}}} \right) \right).$$

### 2.1.4.3 Extreme-value copulas

**Definition 2.1.14** (Extreme-value copulas). A  $d$ -dimensional copula  $C$  is an extreme-value copula if there exists a copula  $C^*$  such that for any  $\mathbf{u} \in [0, 1]^d$ ,

$$\lim_{n \rightarrow \infty} C^*(\mathbf{u}^{1/n})^n = C(\mathbf{u}).$$

The copula  $C^*$  is said to be *in the maximum domain of attraction* of  $C$ .

The following result was proven in Section 3.3 of Hofert et al. (2018).

**Proposition 2.1.15.** *A copula  $C$  is an extreme-value copula if and only if for any  $\mathbf{u} \in [0, 1]^d$  and  $r \in \mathbb{N}$ ,*

$$C(\mathbf{u}^{1/r})^r = C(\mathbf{u}).$$

## 2.2 Estimators of copulas

Let  $(\mathbf{X}_i)_{i \in \mathbb{Z}}$  be a  $d$ -dimensional stationary time series, where the multivariate d.f.  $F$  of  $\mathbf{X}_1$  is assumed to have continuous univariate margins  $F_1, \dots, F_d$ . As a consequence of Theorem 2.1.2, the multivariate d.f.  $F$  can be expressed as

$$F(\mathbf{x}) = C\{F_1(x_1), \dots, F_d(x_d)\}, \quad \mathbf{x} \in \mathbb{R}^d,$$

where the copula  $C$  is unique. Then, suppose we have at hand a sample denoted by  $\mathbf{X}_{1:n} = (\mathbf{X}_1, \dots, \mathbf{X}_n)$  and we wish to estimate  $C$ . First, let us assume that the univariate margins are known and that we are able to define the  $d$ -dimensional stationary time series  $(\mathbf{U}_i)_{i \in \mathbb{Z}}$ , where for  $i \in \mathbb{Z}$ ,

$$\mathbf{U}_i = (F_1(X_{i1}), \dots, F_d(X_{id})). \quad (2.11)$$

Then, according to Lemma 2.1.4, the sample denoted by  $\mathcal{U}_{1:n} = (\mathbf{U}_1, \dots, \mathbf{U}_n)$  is a sample from  $C$  and classical estimation methods could be applied. Unfortunately, the more realistic assumption is that the univariate margins are unknown and we need to estimate them. Based on how the univariate margins and the copula are estimated, we divide the methods of estimation into three categories: parametric estimation, semiparametric estimation and nonparametric estimation.

### 2.2.1 Parametric estimators

In this section, we focus on the estimation of  $C$  under the following two conditions. The notation and examples of this section follows Chapter 4 of Hofert et al. (2018).

**Condition 2.2.1** (Parametric copula family).  $C$  in (2.2) belongs to an absolutely continuous parametric family of copulas

$$\mathcal{C} = \{C_{\boldsymbol{\theta}} : \boldsymbol{\theta} \in \Theta\},$$

where  $\Theta$  is the parameter space and is a subset of  $\mathbb{R}^p$  for some  $p \geq 1$ .

**Condition 2.2.2** (Parametric marginal family).  $F_1, \dots, F_d$  in (2.2) belong to absolutely continuous parametric families of univariate d.f.s  $\mathcal{F}_1, \dots, \mathcal{F}_d$ , respectively, where for any  $j \in \{1, \dots, d\}$ ,

$$\mathcal{F}_j = \{F_{j,\gamma_j} : \gamma_j \in \Gamma_j\},$$

and  $\Gamma_j$  is a subset of  $\mathbb{R}^{p_j}$  for some  $p_j \geq 1$ .

When Conditions 2.2.1 and 2.2.2 are satisfied,  $\boldsymbol{\theta}$  can be estimated with classical parametric estimators, the most famous of which is the maximum likelihood estimator (MLE). Here we provide an example of this method for i.i.d. data.

**Example 2.2.3** (Maximum likelihood estimator in the i.i.d. case). Assume that the random vectors in  $\boldsymbol{\mathcal{X}}_{1:n}$  are i.i.d, and that Conditions 2.2.1 and 2.2.2 are satisfied. Then, according to (2.9), the density  $f$  can be expressed as

$$f(\mathbf{x}) = c_{\boldsymbol{\theta}} \{F_{1,\gamma_1}(x_1), \dots, F_{d,\gamma_d}(x_d)\} \prod_{j=1}^d f_{j,\gamma_j}(x_j),$$

where  $c_{\boldsymbol{\theta}}$  denotes the density of  $C$  and  $f_{j,\gamma_j}$  denotes the density of  $F_{j,\gamma_j}$  for  $j \in \{1, \dots, d\}$ . Then, using the fact that random vectors in  $\boldsymbol{\mathcal{X}}_{1:n}$  are i.i.d, then the maximum likelihood estimator is given by

$$(\hat{\gamma}_{n,1}, \dots, \hat{\gamma}_{n,d}, \hat{\boldsymbol{\theta}}_n) = \underset{(\gamma_1, \dots, \gamma_d, \boldsymbol{\theta}) \in (\Gamma_1, \dots, \Gamma_d, \Theta)}{\operatorname{arg sup}} \ell_n(\gamma_1, \dots, \gamma_d, \boldsymbol{\theta}; \boldsymbol{\mathcal{X}}_{1:n}),$$

where

$$\ell_n(\gamma_1, \dots, \gamma_d, \boldsymbol{\theta}; \boldsymbol{\mathcal{X}}_{1:n}) = \sum_{i=1}^n \log c_{\boldsymbol{\theta}} \{F_{1,\gamma_1}(X_{i1}), \dots, F_{d,\gamma_d}(X_{id})\} + \sum_{j=1}^d \sum_{i=1}^n \log f_{j,\gamma_j}(X_{ij}) \quad (2.12)$$

is the log-likelihood function.

There are, however, two major drawbacks with MLE. Firstly, the estimation of  $\theta$  is impacted by Condition 2.2.1 and 2.2.2. If  $C \notin \mathcal{C}$ , then the estimation of  $C$  is obviously biased. Furthermore, even if  $C \in \mathcal{C}$ , any misspecification of the margins will result in the bias of the estimation of  $\theta$ . Secondly, the log-likelihood function, for instance (2.12) in the i.i.d. case, has to be maximized over a potentially high-dimensional parameter space, and this procedure is known to be difficult in high dimensions. This issue becomes even more significant when  $d$  is large.

One commonly used approach to address the second drawback is called the *inference functions for margins (IFME)* in the literature; see Joe (1997) and Joe and Xu (1996). It separates the estimation procedure into two stages. The first stage is to obtain the marginal estimates  $\hat{\gamma}_{n,1}, \dots, \hat{\gamma}_{n,d}$  and the (*parametric*) *pseudo-observations* (explained in the following example). The second stage is to use the (*parametric*) *pseudo-observations* to obtain the IFME. Here we provide an example of this method for i.i.d. data.

**Example 2.2.4** (Inference Functions for Margins Estimator in the i.i.d. case). Assume that the random vectors in  $\mathcal{X}_{1:n}$  are i.i.d, and that Conditions 2.2.1 and 2.2.2 are satisfied. Then, we could first estimate the unknown marginal parameter vectors by letting, for  $j \in \{1, \dots, d\}$ ,

$$\hat{\gamma}_{n,j} = \arg \sup_{\gamma_j \in \Gamma_j} \sum_{i=1}^n \log f_{j,\gamma_j}(X_{ij}). \quad (2.13)$$

Then we are able to obtain the sample

$$\hat{U}_{i,\hat{\gamma}_n} = (F_{1,\hat{\gamma}_{n,1}}(X_{i1}), \dots, F_{d,\hat{\gamma}_{n,d}}(X_{id})), \quad i \in \{1, \dots, n\}. \quad (2.14)$$

We could refer to  $\hat{U}_{1,\hat{\gamma}_n}, \dots, \hat{U}_{n,\hat{\gamma}_n}$  as (*parametric*) *pseudo-observations* from  $C$ . Finally, we are able to estimate  $\theta$  by letting

$$\hat{\theta}_n = \arg \sup_{\theta \in \Theta} \sum_{i=1}^n \log c_{\theta}(\hat{U}_{i,\hat{\gamma}_n}). \quad (2.15)$$

Basically, the IFME separates the maximization problem over a high-dimensional parameter space into a maximization problem of  $d+1$  lower-dimensional parameter spaces. In this way, it greatly reduces the computational difficulty, although, according to Joe (2005), suffers from an efficiency loss with respect to the full MLE.

## 2.2.2 Semiparametric estimators

To overcome the issue of the misspecification of the margins, we could estimate the marginal d.f.s nonparametrically. Specifically, for  $j \in \{1, \dots, d\}$ , let  $F_{1:n,j}$  be the empirical d.f. of the  $j$ th

component sample  $X_{1j}, \dots, X_{nj}$  defined by

$$F_{1:n,j}(x) = \frac{1}{n} \sum_{i=1}^n \mathbf{1}\{X_{ij} \leq x\}, \quad x \in \mathbb{R}. \quad (2.16)$$

Then, starting with (2.11) and replacing  $F_j$  with  $F_{1:n,j}$ , we are able to define the observable random vectors

$$\hat{U}_i^{1:n} = \mathbf{R}_i^{1:n}/n = (F_{1:n,1}(X_{i1}), \dots, F_{1:n,d}(X_{id})), \quad i \in \{1, \dots, n\}, \quad (2.17)$$

where for any  $i \in \{1, \dots, n\}$ ,  $\mathbf{R}_i^{1:n} = (R_{i1}^{1:n}, \dots, R_{id}^{1:n})$  with  $R_{ij}^{1:n} = nF_{1:n,j}(X_{ij}) = \sum_{t=1}^n \mathbf{1}(X_{tj} \leq X_{ij})$  being the (maximal) rank of  $X_{ij}$  among  $X_{1j}, \dots, X_{nj}$ . Note that  $\mathbf{R}_i^{1:n}$  and  $\hat{U}_i^{1:n}$  are the multivariate ranks and the multivariate scaled ranks, respectively, obtained from  $\mathcal{X}_{1:n}$ . Analogous to  $\hat{U}_{1,\hat{\gamma}_n}, \dots, \hat{U}_{n,\hat{\gamma}_n}$  in (2.14),  $\hat{U}_1^{1:n}, \dots, \hat{U}_n^{1:n}$  can be referred to as (*nonparametric*) *pseudo-observations* from  $C$ . Then, similar to IFME, the *maximum pseudo-likelihood estimator*, can be viewed as another two-stage estimator: the first stage is to estimate the margins nonparametrically, and then the second stage is to estimate  $\theta$  with the (nonparametric) pseudo-observations; see Genest et al. (1995) and Klassen and Wellner (1997). The asymptotic distribution of this estimator was studied by Genest et al. (1995) and Shih and Louis (1995) in the i.i.d. case and by Chen and Fan (2006) in the time series case. Here we provide an example of this method for i.i.d. data.

**Example 2.2.5** (Maximum pseudo-likelihood estimators in the i.i.d. case). Assume that random vectors in  $\mathcal{X}_{1:n}$  are i.i.d, and that Condition 2.2.1 is satisfied. Then, starting from (2.15) and replacing  $\hat{U}_{i,\hat{\gamma}_n}$  by  $n\hat{U}_i^{1:n}/(n+1)$ , we obtain the maximum pseudo-likelihood estimator defined by

$$\hat{\theta} = \arg \sup_{\theta \in \Theta} \sum_{i=1}^n \log c_{\theta} \left( \frac{n}{n+1} \hat{U}_i^{1:n} \right). \quad (2.18)$$

Note that the reason to include the rescaling term  $n/(n+1)$  is to ensure that  $n\hat{U}_i^{1:n}/(n+1)$  lies in the interior  $(0,1)^d$  of the unit hypercube since several copula families have non-finite density on the edges of  $[0,1]^d$  ending at  $(1, \dots, 1)$ . Nevertheless, one can show that the rescaling term is asymptotically negligible (see, e.g., Kojadinovic et al., 2011; Hofert et al., 2018).

### 2.2.3 Nonparametric estimators

In the previous section, although the semiparametric estimation overcomes the problem of misspecification of the margins, the family of copula could still be misspecified. In addition, complex dependence structure may still be hard to capture within the parametric framework. Therefore, a fully nonparametric estimator that relies neither on Condition 2.2.1 nor 2.2.2 could be advocated.



### 2.2.3.1 The empirical copula

Let us first define the sample  $\mathcal{U}_{1:n} = (\mathbf{U}_1, \dots, \mathbf{U}_n)$ , where  $\mathbf{U}_i$  is defined in (2.11) and has d.f.  $C$ . Therefore a natural but unobservable empirical copula can be defined by

$$G_{1:n}(\mathbf{u}) = \frac{1}{n} \sum_{i=1}^n \mathbf{1}(\mathbf{U}_i \leq \mathbf{u}), \quad \mathbf{u} \in [0, 1]^d, \quad (2.19)$$

where inequalities between vectors are understood componentwise. Unfortunately  $G_{1:n}$  cannot be computed since  $\mathbf{U}_i$  is unobservable and one way to remedy this is to replace  $\mathbf{U}_i$  by the nonparametric pseudo-observations  $\hat{\mathbf{U}}_i^{1:n}$  defined in (2.17). This leads to the frequently used version of the empirical copula (see, e.g., Genest et al., 1995) defined by

$$C_{1:n}(\mathbf{u}) = \frac{1}{n} \sum_{i=1}^n \mathbf{1}(\hat{\mathbf{U}}_i^{1:n} \leq \mathbf{u}), \quad \mathbf{u} \in [0, 1]^d. \quad (2.20)$$

Notice that another version of the empirical copula could also be considered. Let  $F_{1:n}$  be the empirical d.f. of  $\mathcal{X}_{1:n}$  defined by

$$F_{1:n}(\mathbf{x}) = \frac{1}{n} \sum_{i=1}^n \mathbf{1}(\mathbf{X}_i \leq \mathbf{x}), \quad \mathbf{x} \in \mathbb{R}^d, \quad (2.21)$$

Then, starting from (2.3) and replacing  $F$  (resp.  $F_j$ ) with  $F_{1:n}$  (resp.  $F_{1:n,j}$ ) leads to another version of the empirical copula (see, e.g., Deheuvels, 1979) defined by

$$\hat{C}_{1:n}(\mathbf{u}) = F_{1:n}\{F_{1:n,1}^{-1}(u_1), \dots, F_{1:n,d}^{-1}(u_d)\}, \quad \mathbf{u} \in [0, 1]^d, \quad (2.22)$$

where, for  $j \in \{1, \dots, d\}$ ,  $F_{1:n,j}^{-1}$  is the generalized inverse of  $F_{1:n,j}$  in (2.16).

**Condition 2.2.6** (No ties). With probability 1, there are no ties in each of the component samples  $X_{1j}, \dots, X_{nj}$ ,  $j \in \{1, \dots, d\}$ , of  $\mathcal{X}_{1:n}$ .

**Lemma 2.2.7.** Under Condition 2.2.6, with probability 1,

$$\sup_{\mathbf{u} \in [0, 1]^d} |\hat{C}_{1:n}(\mathbf{u}) - C_{1:n}(\mathbf{u})| \leq \frac{d}{n}.$$

*Proof.* We proceed as in the proof of Proposition 3.3 in Bücher et al. (2014). Fix  $\mathbf{u} \in [0, 1]^d$  and  $n \in \mathbb{N}$ . For  $i \in \{1, \dots, n\}$  and  $j \in \{1, \dots, d\}$ , using the fact that, for any d.f.  $H$ ,  $x < H^{-1}(u)$  if and only if  $H(x) < u$ , we have that

$$\mathbf{1}\{X_{ij} \leq F_{1:n,j}^{-1}(u_j)\} - \mathbf{1}\{F_{1:n,j}(X_{ij}) \leq u_j\} = \mathbf{1}\{X_{ij} = F_{1:n,j}^{-1}(u_j)\} - \mathbf{1}\{F_{1:n,j}(X_{ij}) = u_j\}. \quad (2.23)$$

Then, using additionally the fact that  $F_{1:n,j}(X_{ij}) = u_j$  implies that  $X_{ij} = F_{1:n,j}^{-1}(u_j)$ ; see also proof of Lemma 7.2 in [Bücher and Segers \(2014\)](#), we have

$$0 \leq \mathbf{1} \left\{ X_{ij} = F_{1:n,j}^{-1}(u_j) \right\} - \mathbf{1} \left\{ F_{1:n,j}(X_{ij}) = u_j \right\} \leq \mathbf{1} \left\{ X_{ij} = F_{1:n,j}^{-1}(u_j) \right\} \quad (2.24)$$

Furthermore, we notice that if  $a_i, b_i \in [0, 1]$  for  $i \in \{1, \dots, d\}$ , then

$$\left| \prod_{i=1}^d a_i - \prod_{i=1}^d b_i \right| = \left| \sum_{i=1}^d (a_i - b_i) \prod_{j=1}^{i-1} b_j \prod_{k=i+1}^d a_k \right| \leq \sum_{i=1}^d |a_i - b_i|. \quad (2.25)$$

Having (2.23), (3.38) and (2.25) in mind and starting from (2.20) and (2.22), we have that

$$\begin{aligned} |\hat{C}_{1:n}(\mathbf{u}) - C_{1:n}(\mathbf{u})| &= \left| \frac{1}{n} \sum_{i=1}^n \prod_{j=1}^d \mathbf{1} \left\{ X_{ij} \leq F_{1:n,j}^{-1}(u_j) \right\} - \frac{1}{n} \sum_{i=1}^n \prod_{j=1}^d \mathbf{1} \left\{ F_{1:n,j}(X_{ij}) \leq u_j \right\} \right| \\ &\leq \frac{1}{n} \sum_{i=1}^n \left| \prod_{j=1}^d \mathbf{1} \left\{ X_{ij} \leq F_{1:n,j}^{-1}(u_j) \right\} - \prod_{j=1}^d \mathbf{1} \left\{ F_{1:n,j}(X_{ij}) \leq u_j \right\} \right| \\ &\leq \frac{1}{n} \sum_{i=1}^n \sum_{j=1}^d \left| \mathbf{1} \left\{ X_{ij} \leq F_{1:n,j}^{-1}(u_j) \right\} - \mathbf{1} \left\{ F_{1:n,j}(X_{ij}) \leq u_j \right\} \right| \\ &= \frac{1}{n} \sum_{i=1}^n \sum_{j=1}^d \mathbf{1} \left\{ X_{ij} = F_{1:n,j}^{-1}(u_j) \right\} = \frac{d}{n}, \end{aligned}$$

where the last equality follows from the fact that for any  $j \in \{1, \dots, d\}$  and  $u_j \in [0, 1]$ ,  $\sum_{i=1}^n \mathbf{1} \left\{ X_{ij} = F_{1:n,j}^{-1}(u_j) \right\} = 1$  since  $F_{1:n,j}^{-1}(u_j)$  is a fixed number that takes value in the set  $\{X_{1j}, \dots, X_{nj}\}$ ; see also proof of Lemma 4.6 in [Berghaus et al. \(2017\)](#). ■

### 2.2.3.2 The empirical checkerboard copula

It can be easily verified that, whether Condition 2.2.6 is satisfied or not, both  $C_{1:n}$  in (2.20) and  $\hat{C}_{1:n}$  in (2.22) are step functions, whereas, a genuine copula is a Lipschitz continuous function according to Proposition 2.1.6. Therefore, certain smoothing methods could be applied to reduce the estimation bias. One smooth estimator is called the *empirical checkerboard copula*, which can be regarded as a multilinear extension of  $C_{1:n}$  in (2.20) (see, e.g., [Genest et al., 2017](#); [Genest and Nešlehová, 2007](#), and the references therein). Specifically, under Condition 2.2.6, the empirical checkerboard copula of  $\mathcal{X}_{1:n}$  is defined by

$$C_{1:n}^{check}(\mathbf{u}) = \frac{1}{n} \sum_{i=1}^n \prod_{j=1}^d \min\{\max\{nu_j - R_{ij}^{1:n} + 1, 0\}, 1\}, \quad \mathbf{u} \in [0, 1]^d, \quad (2.26)$$

and is a genuine copula.

### 2.2.3.3 The empirical Bernstein copula and the empirical beta copula

The *empirical Bernstein copula* was first proposed in [Sancetta and Satchell \(2004\)](#) and later studied by [Segers et al. \(2017\)](#). Specifically, let  $\mathbf{m} = (m_1, \dots, m_d) \in \mathbb{N}^d$  and define

$$C_{1:n,\mathbf{m}}^{\text{Bern}}(\mathbf{u}) = \int_{[0,1]^d} C_{1:n}(\mathbf{w}) d\mu_{\mathbf{m},\mathbf{u}}(\mathbf{w}), \quad \mathbf{u} \in [0,1]^d, \quad (2.27)$$

where  $\mu_{\mathbf{m},\mathbf{u}}$  is the law of the  $d$ -dimensional random vector  $(S_{m_1,1,u_1}/n, \dots, S_{m_d,d,u_d}/n)$ , where  $S_{m_1,1,u_1}, \dots, S_{m_d,d,u_d}$  are independent random variables and, for each  $j \in \{1, \dots, d\}$ ,  $S_{m_j,j,u_j}$  is Binomial( $m_j, u_j$ ), that is,  $S_{m_j,j,u_j}$  follows a binomial distribution with parameters  $m_j$  and  $u_j$ . Building upon the work of [Sancetta and Satchell \(2004\)](#) and checking the three conditions of [Theorem 2.1.5](#), [Segers et al. \(2017\)](#) showed that  $C_{1:n,\mathbf{m}}^{\text{Bern}}$  is a genuine copula if and only if all the polynomial degrees  $m_1, \dots, m_d$  are divisors of  $n$ . With this idea in mind, [Segers et al. \(2017\)](#) proposed the *empirical beta copula* as a special case of the empirical Bernstein copula by letting  $\mathbf{m} = (n, \dots, n)$ . Specifically, for  $n \in \mathbb{N}$ ,

$$C_{1:n}^{\text{beta}}(\mathbf{u}) = \int_{[0,1]^d} C_{1:n}(\mathbf{w}) d\mu_{n,\mathbf{u}}(\mathbf{w}), \quad (2.28)$$

where  $\mu_{n,\mathbf{u}}$  corresponds to  $\mu_{\mathbf{m},\mathbf{u}}$  when  $\mathbf{m} = (n, \dots, n)$ . The name “beta” comes from the fact that it can be expressed in the following form:

$$C_{1:n}^{\text{beta}}(\mathbf{u}) = \frac{1}{n} \sum_{i=1}^n \prod_{j=1}^d \text{beta}_{n,R_{ij}^{1:n}}(u_j), \quad \mathbf{u} \in [0,1]^d, \quad (2.29)$$

where, for any  $n \in \mathbb{N}$  and  $r \in \{1, \dots, n\}$ ,  $\text{beta}_{n,r}$  denotes the d.f. of the distribution Beta( $r, n + 1 - r$ ) (the beta distribution with shape parameters  $\alpha = r$  and  $\beta = n + 1 - r$ ).

**Proposition 2.2.8** ([Segers et al. \(2017\)](#)). *Under Condition 2.2.6, we have*

$$\sup_{\mathbf{u} \in [0,1]^d} \left| C_{1:n}(\mathbf{u}) - C_{1:n}^{\text{beta}}(\mathbf{u}) \right| \leq d \{n^{-1/2}(\log n)^{1/2} + n^{-1/2} + n^{-1}\}.$$

## 2.3 Empirical copula processes

The empirical copula  $C_{1:n}$  in (2.20) has been applied in many nonparametric inference procedures on  $C$ . For instance, it was used in [Genest and Rémillard \(2004\)](#) for tests of independence, in [Rémillard and Scaillet \(2009\)](#) for tests of equality between two copulas, and in [Genest et al. \(2012\)](#) for tests of symmetry for bivariate copulas. The asymptotics of the aforementioned procedures usually follows from the asymptotics of the *empirical copula process*. The latter has been studied extensively, among others, by [Rüschendorf \(1976\)](#), [Gänssler and Stute \(1987\)](#), [Fermanian et al. \(2004\)](#), [Tsukahara \(2005\)](#), [van der Vaart and Wellner \(2007\)](#), [Segers \(2012\)](#) and [Bücher and Volgushev \(2013\)](#) among others. Furthermore, the *empirical Bernstein copula*

process was proposed and studied in Segers et al. (2017). In addition, in the study of *change point detection* (see, e.g., Bücher et al., 2014), a more general *two-sided empirical copula process* was proposed and studied in Bücher and Kojadinovic (2016).

In this section, we first introduce the mathematical preliminaries needed to study the asymptotics of all the aforementioned processes. Then, we present some important asymptotic results in the recent literature. Furthermore, unless stated otherwise, all convergences in this work are with respect to  $n \rightarrow \infty$ .

### 2.3.1 Mathematical preliminaries

Let us recall the definition of weak convergence of random vectors in Lemma 2.2 in van der Vaart (1998). We say that a sequence of random elements  $\mathbf{X}_n$  in  $\mathbb{R}^d$  converges in distribution to  $\mathbf{X}$  if and only if,

$$\mathbb{E}f(\mathbf{X}_n) \rightarrow \mathbb{E}f(\mathbf{X}), \quad \text{for all } f \in C_b(\mathbb{R}^d), \quad (2.30)$$

where  $C_b(\mathbb{R}^d)$  denotes the set of all continuous and bounded functions on  $\mathbb{R}^d$ . As a natural generalization, we would like to say that a sequence of random elements  $\mathbb{X}_n$  in arbitrary metric space  $\mathbb{D}$  converges in distribution to an element  $\mathbb{X} \in \mathbb{D}$  if and only if,

$$\mathbb{E}f(\mathbb{X}_n) \rightarrow \mathbb{E}f(\mathbb{X}), \quad \text{for all } f \in C_b(\mathbb{D}), \quad (2.31)$$

where  $C_b(\mathbb{D})$  denotes the set of all continuous and bounded functions on  $\mathbb{D}$ . Unfortunately, the expression in (2.31) is, in general, wrong since measurability issues might easily occur when  $\mathbb{D}$  is not separable. The next example shows how measurability issues occur when  $\mathbb{D} = D([0, 1])$ , the space of *càdlàg functions* (right-continuous with left limits) equipped with the metric induced by the supremum norm.

**Example 2.3.1.** Let  $U$  be a standard uniform random variable defined on a probability space  $(\Omega, \mathcal{A}, \mathbb{P})$ . For every  $\omega \in \Omega$ , define  $\mathbb{X}(t)(\omega) = \mathbf{1}_{[U(\omega), 1]}(t)$ ,  $t \in [0, 1]$ . Let  $\mathcal{B}(D[0, 1])$  be the Borel  $\sigma$ -field generated by the open sets in  $D([0, 1])$ . Then for any  $s \in [0, 1]$ ,  $\mathbf{1}_{[s, 1]}(\cdot)$  is an element of  $D([0, 1])$ . Next, let us define the open ball in  $D([0, 1])$  with center  $\mathbf{1}_{[s, 1]}$  with radius  $1/2$  by

$$B_{1/2}(\mathbf{1}_{[s, 1]}) = \left\{ f \in D([0, 1]) : \sup_{t \in [0, 1]} |f(t) - \mathbf{1}_{[s, 1]}(t)| < 1/2 \right\}.$$

Let  $V$  be a non-measurable set on  $[0, 1]$  and let  $W = \cup_{s \in V} B_{1/2}(\mathbf{1}_{[s, 1]})$ . We can easily verify that  $W \in \mathcal{B}(D[0, 1])$  and that  $\{\omega : \mathbb{X}(\omega) \in W\} = \{\omega : U(\omega) \in V\}$ . However, since  $\{\omega : U(\omega) \in V\} \notin \mathcal{A}$ , we conclude that  $\mathbb{X}$  is not  $\mathcal{A}/\mathcal{B}(D[0, 1])$ -measurable.

During the last several decades, several approaches to overcome this difficulty were suggested. In this section, we will briefly summarize the most modern one, studied extensively in van der Vaart and Wellner (2000) and Kosorok (2008). The key idea is to drop the requirement of Borel

measurability of each  $\mathbb{X}_n$ , meanwhile keeping the requirement (2.31), where expectations are replaced by outer expectations.

**Definition 2.3.2** (Outer expectation and outer probability). Let  $(\Omega, \mathcal{A}, \mathbb{P})$  be an arbitrary probability space and  $X : \Omega \mapsto \bar{\mathbb{R}}$  be an arbitrary map, where  $\bar{\mathbb{R}}$  is the extended real line. The outer expectation of  $X$  with respect to  $\mathbb{P}$  is defined as

$$\mathbb{E}^* X = \inf \{ \mathbb{E} U : U \geq X, U; \Omega \mapsto \bar{\mathbb{R}} \text{ measurable and } \mathbb{E} X \text{ exists} \}.$$

The outer probability of an arbitrary subset  $B \subset \Omega$  is

$$\mathbb{P}^*(B) = \inf \{ P(A) : A \supset B, A \in \mathcal{A} \}.$$

*Inner expectation* and *inner probability* can be defined by  $\mathbb{E}_* X = -\mathbb{E}^*(-X)$  and  $\mathbb{P}_*(B) = 1 - \mathbb{P}^*(\Omega - B)$ , respectively.

**Definition 2.3.3** (Weak convergence). Let  $(\Omega_n, \mathcal{A}_n, \mathbb{P}_n)$  be a sequence of probability spaces and let  $\mathbb{X}_n : \Omega_n \mapsto \mathbb{D}$  be arbitrary maps. Furthermore, let  $\mathbb{X} : \Omega \mapsto \mathbb{D}$  be a Borel measurable map defined on  $(\Omega, \mathcal{A}, \mathbb{P})$ . Then we say that  $\mathbb{X}_n$  converges weakly to  $\mathbb{X}$ , and write  $\mathbb{X}_n \rightsquigarrow \mathbb{X}$  in  $\mathbb{D}$ , if and only if

$$\mathbb{E}^* f(\mathbb{X}_n) \rightarrow \mathbb{E} f(\mathbb{X}), \quad \text{for all } f \in C_b(\mathbb{D}).$$

In the rest of this section, we consider the case  $\mathbb{D} = \ell^\infty(T)$ , where  $\ell^\infty(T)$  is the space of all bounded real-valued functions on  $T$  equipped with the uniform metric.

**Definition 2.3.4** (Asymptotically measurable). The sequence of maps  $\mathbb{X}_n$  is asymptotically measurable if and only if

$$\mathbb{E}^* f(\mathbb{X}_n) - \mathbb{E}_* f(\mathbb{X}_n) \rightarrow 0, \quad \text{for every } f \in C_b(\mathbb{D}).$$

**Definition 2.3.5** (Asymptotic tightness). The sequence  $\mathbb{X}_n$  is asymptotically tight if and only if for every  $\varepsilon > 0$  there exists a compact set  $K$  such that

$$\liminf \mathbb{P}_*(\mathbb{X}_n \in K^\delta) \geq 1 - \varepsilon, \quad \text{for every } \delta > 0.$$

Here  $K^\delta = \{y \in \mathbb{D} : d(y, K) \leq \delta\}$  is the “ $\delta$ -enlargement” around  $K$ .

The proof of the following result can be found in Theorem 1.5.6 and Theorem 1.5.7 in [van der Vaart and Wellner \(2000\)](#).

**Theorem 2.3.6.** *Suppose  $\mathbb{X}_n : \Omega_n \mapsto \ell^\infty(T)$  is a sequence of maps. Then the following three statements are equivalent:*

- (i) *The sequence  $\mathbb{X}_n$  is asymptotically tight.*

(ii) For all  $\varepsilon, \eta$ , there exists a partition  $T = \cup_{i=1}^k T_i$  such that

$$\limsup_{n \rightarrow \infty} \mathbb{P}^* \left( \sup_i \sup_{s, t \in T_i} |\mathbb{X}_n(s) - \mathbb{X}_n(t)| > \varepsilon \right) < \eta.$$

(iii) For every  $t \in T$ , there exists a semimetric  $\rho$  on  $T$  such that  $(T, \rho)$  is totally bounded and for every  $\varepsilon, \eta > 0$ , there exists a  $\delta > 0$  such that

$$\limsup_{n \rightarrow \infty} \mathbb{P}^* \left( \sup_{\rho(s, t) < \delta} |\mathbb{X}_n(s) - \mathbb{X}_n(t)| > \varepsilon \right) < \eta.$$

The proof of the following result can be found in Theorem 1.5.4 in [van der Vaart and Wellner \(2000\)](#).

**Theorem 2.3.7** (Weak convergence). *Let  $(\Omega_n, \mathcal{A}_n, \mathbb{P}_n)$  be a sequence of probability spaces and let  $\mathbb{X}_n : \Omega_n \mapsto \ell^\infty(T)$  be arbitrary maps. Furthermore, let  $\mathbb{X} : \Omega \mapsto \ell^\infty(T)$  be a Borel measurable map defined on  $(\Omega, \mathcal{A}, \mathbb{P})$ . Then we say that  $\mathbb{X}_n \rightsquigarrow \mathbb{X}$  in  $\ell^\infty(T)$ , if and only if*

(i) *The sequence  $\mathbb{X}_n$  is asymptotically tight.*

(ii) *For every finite subset  $t_1, \dots, t_k$  of  $T$ , the random vector  $(\mathbb{X}_n(t_1), \dots, \mathbb{X}_n(t_k))$  converges weakly to  $(\mathbb{X}(t_1), \dots, \mathbb{X}(t_k))$  in  $\mathbb{R}^k$ .*

### 2.3.2 Empirical copula processes

First, let us explain how the quantities introduced in Section 2.2.3.1 can be defined from sub-stretches of the available stretch of observations  $\mathbf{X}_{1:n} = (\mathbf{X}_1, \dots, \mathbf{X}_n)$ . Given a substretch  $\mathbf{X}_{k:l} = (\mathbf{X}_k, \dots, \mathbf{X}_l)$ ,  $1 \leq k \leq l \leq n$ , let  $F_{k:l}$  be the empirical d.f. of  $\mathbf{X}_{k:l}$  defined by

$$F_{k:l}(\mathbf{x}) = \frac{1}{l - k + 1} \sum_{i=k}^l \mathbf{1}\{\mathbf{X}_i \leq \mathbf{x}\}, \quad \mathbf{x} \in \mathbb{R}^d, \quad (2.32)$$

and for  $j \in \{1, \dots, d\}$ , let  $F_{k:l,j}$  be the empirical d.f.s of the  $j$ th component sample  $X_{kj}, \dots, X_{lj}$  defined by

$$F_{k:l,j}(x_j) = \frac{1}{l - k + 1} \sum_{i=k}^l \mathbf{1}\{X_{ij} \leq x_j\}, \quad x \in \mathbb{R}. \quad (2.33)$$

Then, we are able to define the observable random vectors

$$\hat{U}_i^{k:l} = \mathbf{R}_i^{k:l} / n = (F_{k:l,1}(X_{i1}), \dots, F_{k:l,d}(X_{id})), \quad i \in \{1, \dots, n\}, \quad (2.34)$$

where for any  $i \in \{1, \dots, n\}$ ,  $\mathbf{R}_i^{k:l} = (R_{i1}^{k:l}, \dots, R_{id}^{k:l})$  and with  $R_{ij}^{k:l} = (l - k + 1)F_{k:l,j}(X_{ij}) = \sum_{t=k}^l \mathbf{1}(X_{tj} \leq X_{ij})$  being the (maximal) rank of  $X_{ij}$  among  $X_{kj}, \dots, X_{lj}$ . Then, for any  $1 \leq$

$k \leq l \leq n$ , we define

$$C_{k:l}(\mathbf{u}) = \frac{1}{l-k+1} \sum_{i=k}^l \mathbf{1}(\hat{U}_i^{k:l} \leq \mathbf{u}), \quad \mathbf{u} \in [0, 1]^d, \quad (2.35)$$

$$\hat{C}_{k:l}(\mathbf{u}) = F_{k:l} \left( F_{k:l,1}^{-1}(u_1), \dots, F_{k:l,d}^{-1}(u_d) \right), \quad \mathbf{u} \in [0, 1]^d, \quad (2.36)$$

where, for  $j \in \{1, \dots, d\}$ ,  $F_{k:l,j}^{-1}$  is the generalized inverse of  $F_{k:l,j}$  in (2.33). Note that, if  $k > l$ , we adopt the convention that  $C_{k:l} = \hat{C}_{k:l} = 0$ . Following Bücher and Kojadinovic (2016), we let  $\Delta = \{(s, t) \in [0, 1]^2 : s \leq t\}$  and let  $\lambda_n(s, t) = (\lfloor nt \rfloor - \lfloor ns \rfloor)/n$ ,  $(s, t) \in \Delta$ . Then the two versions of the two-sided empirical copula process are defined by

$$\mathbb{C}_n(s, t, \mathbf{u}) = \sqrt{n} \lambda_n(s, t) \{C_{\lfloor ns \rfloor + 1 : \lfloor nt \rfloor}(\mathbf{u}) - C(\mathbf{u})\}, \quad (s, t, \mathbf{u}) \in \Delta \times [0, 1]^d, \quad (2.37)$$

$$\hat{\mathbb{C}}_n(s, t, \mathbf{u}) = \sqrt{n} \lambda_n(s, t) \{\hat{C}_{\lfloor ns \rfloor + 1 : \lfloor nt \rfloor}(\mathbf{u}) - C(\mathbf{u})\}, \quad (s, t, \mathbf{u}) \in \Delta \times [0, 1]^d. \quad (2.38)$$

Combined with (2.35),  $\mathbb{C}_n$  can also be written equivalently as

$$\mathbb{C}_n(s, t, \mathbf{u}) = \frac{1}{\sqrt{n}} \sum_{i=\lfloor ns \rfloor + 1}^{\lfloor nt \rfloor} \left\{ \mathbf{1}(\hat{U}_i^{\lfloor ns \rfloor + 1 : \lfloor nt \rfloor} \leq \mathbf{u}) - C(\mathbf{u}) \right\}, \quad (s, t, \mathbf{u}) \in \Delta \times [0, 1]^d.$$

### 2.3.2.1 Asymptotics in the i.i.d. case

When the underlying sequence of random vectors  $(\mathbf{X}_i)_{i \in \mathbb{Z}}$  are i.i.d, the asymptotics of  $\mathbb{C}_n(0, 1, \cdot)$  in (2.37) have been studied, among others, in Rüschendorf (1976), Gänssler and Stute (1987), van der Vaart and Wellner (2000, Chapter 3, page 389), Fermanian et al. (2004) and Tsukahara (2005) under the assumption that  $C$  has continuous partial derivatives on  $[0, 1]^d$ . In a later study, Segers (2012) pointed out that the latter assumption is too strong for many popular copula families and showed that the results can be proven under the following weaker condition.

**Condition 2.3.8** (Smooth partial derivatives). For any  $j \in \{1, \dots, d\}$ , the partial derivative  $\dot{C}_j = \partial C / \partial u_j$  exists and is continuous on the set  $V_j = \{\mathbf{u} \in [0, 1]^d : u_j \in (0, 1)\}$ .

*Remark 2.3.9.* It is explained in Segers (2012) that Condition 2.3.8 is satisfied in the case of Archimedean and extreme-value copulas (defined in Definition 2.1.9 and 2.1.14 respectively). However, Condition 2.3.8 does not hold, for instance, for the Fréchet-Hoeffding lower (resp. upper) bounds, where  $C(u_1, u_2) = \max(u_1 + u_2 - 1, 0)$  (resp.  $C(u_1, u_2) = \min(u_1, u_2)$ ).

**Theorem 2.3.10.** Segers (2012) Suppose that the sequence  $(\mathbf{X}_i)_{i \in \mathbb{Z}}$  is i.i.d. and that Condition 2.3.8 is satisfied. Then we have that  $\mathbb{C}_n(0, 1, \cdot) \rightsquigarrow \mathbb{G}_C$  in  $\ell^\infty([0, 1]^d)$ , where

$$\mathbb{G}_C(\mathbf{u}) = \mathbb{U}_C(\mathbf{u}) - \sum_{j=1}^d \dot{C}_j(\mathbf{u}) \mathbb{U}_C(\mathbf{u}^{(j)}), \quad \mathbf{u} \in [0, 1]^d, \quad (2.39)$$

and  $\mathbb{U}_C$  is a tight centered Gaussian process with covariance function

$$\text{Cov}(\mathbb{U}_C(\mathbf{u}), \mathbb{U}_C(\mathbf{v})) = C(\mathbf{u} \wedge \mathbf{v}) - C(\mathbf{u})C(\mathbf{v}).$$

### 2.3.2.2 Asymptotics in the time series case

In a further step, [Bücher and Volgushev \(2013\)](#) extended the results in [Segers \(2012\)](#) to the setting where the underlying stationary time series  $(\mathbf{X}_i)_{i \in \mathbb{Z}}$  is serially dependent. Let  $\Phi$  be the map from  $\ell^\infty([0, 1]^d)$  to  $\ell^\infty([0, 1]^d)$  defined for any d.f.  $H$  on  $[0, 1]^d$  whose univariate margins  $H_1, \dots, H_d$  do not assign mass at zero by

$$\Phi(H)(\mathbf{u}) = H\{H_1^{-1}(u_1), \dots, H_d^{-1}(u_d)\}, \quad \mathbf{u} \in [0, 1]^d. \quad (2.40)$$

The following result was proven in Theorem 2.4 of [Bücher and Volgushev \(2013\)](#).

**Theorem 2.3.11** ([Bücher and Volgushev \(2013\)](#)). *Suppose Condition 2.3.8 holds. Then  $\Phi$  in (2.40) is Hardamard-differentiable at  $C$  tangentially to  $\mathcal{D}_0$ , where*

$$\mathcal{D}_0 = \{f \in \mathcal{C}([0, 1]^d) : f(\mathbf{u}) = 0 \text{ if one of the components of } \mathbf{u} \text{ is } 0, \text{ and } f(1, \dots, 1) = 0\}. \quad (2.41)$$

Its derivative at  $C$  in  $\alpha \in \mathcal{D}_0$  is given by

$$\Phi'_C(\alpha)(\mathbf{u}) = \alpha(\mathbf{u}) - \sum_{j=1}^d C_j(\mathbf{u})\alpha(\mathbf{u}^{(j)}).$$

Now, let us first define a two-sided sequential empirical process

$$\mathbb{B}_n(s, t, \mathbf{u}) = \frac{1}{\sqrt{n}} \sum_{i=[ns]+1}^{[nt]} \{\mathbf{1}(\mathbf{U}_i \leq \mathbf{u}) - C(\mathbf{u})\}, \quad (s, t, \mathbf{u}) \in \Delta \times [0, 1]^d, \quad (2.42)$$

where  $\mathbf{U}_1, \dots, \mathbf{U}_n$  are defined in (2.11), and with the convention that  $\mathbb{B}_n(s, t, \cdot) = 0$  if  $[nt] - [ns] = 0$ .

Then, for the convergence of  $\mathbb{C}_n(0, 1, \cdot)$  to hold, [Bücher and Volgushev \(2013\)](#) found that it suffices to impose the following condition.

**Condition 2.3.12** (Weak convergence of  $\mathbb{B}_n(0, 1, \cdot)$  in the time series case). The sequential empirical process  $\mathbb{B}_n(0, 1, \cdot)$  in (2.42) converges weakly in  $\ell^\infty([0, 1]^d)$  to a tight centered Gaussian process  $\mathbb{Z}_C(1, \cdot)$  concentrated on  $\mathcal{D}_0$  defined in (2.41).

Note that Condition 2.3.12 is a standard result of empirical process theory for weakly dependent data, where the limiting process  $\mathbb{Z}_C(1, \cdot)$  is a tight, centered Gaussian process on  $[0, 1]^d$  with



covariance function

$$\text{Cov}(\mathbb{Z}_C(\mathbf{1}, \mathbf{u}), \mathbb{Z}_C(\mathbf{1}, \mathbf{v})) = \sum_{k \in \mathbb{Z}} \text{Cov}\{\mathbf{1}(U_0 \leq \mathbf{u}), \mathbf{1}(U_k \leq \mathbf{v})\}.$$

Then, according to [Bücher and Volgushev \(2013\)](#), the following result is an immediate consequence of [Theorem 2.3.11](#) and the functional delta method ([van der Vaart and Wellner, 2000](#), [Theorem 3.9.4](#), page 375).

**Theorem 2.3.13** ([Bücher and Volgushev \(2013\)](#)). *Suppose that the sequence  $(\mathbf{X}_i)_{i \in \mathbb{Z}}$  is a stationary time series that satisfies [Condition 2.3.8](#) and [2.3.12](#). Then we have that,  $\hat{\mathbb{C}}_n(0, 1, \cdot) \rightsquigarrow \mathbb{C}_C(0, 1, \cdot)$  in  $\ell^\infty([0, 1]^d)$ , where  $\hat{\mathbb{C}}(0, 1, \cdot)$  is defined in [\(2.38\)](#) and*

$$\mathbb{C}_C(0, 1, \mathbf{u}) = \mathbb{Z}_C(\mathbf{1}, \mathbf{u}) - \sum_{j=1}^d \dot{C}_j(\mathbf{u}) \mathbb{Z}_C(\mathbf{1}, \mathbf{u}^{(j)}), \quad \mathbf{u} \in [0, 1]^d.$$

*Remark 2.3.14.* Note that, under [Condition 2.2.6](#), we have that, according to [Lemma 2.2.7](#),  $\sup_{\mathbf{u} \in [0, 1]^d} |\hat{\mathbb{C}}_n(0, 1, \cdot) - \mathbb{C}_n(0, 1, \cdot)| \leq d/\sqrt{n}$ . In this case, the asymptotics of  $\mathbb{C}_n(0, 1, \cdot)$  in [\(2.37\)](#) and  $\hat{\mathbb{C}}_n(0, 1, \cdot)$  in [\(2.38\)](#) are the same.

In particular, the following *strongly mixing* condition entails [Condition 2.3.12](#) (see, e.g., [Philipp and Pinzur, 1980](#), [Theorem 2](#)).

**Definition 2.3.15** (Strongly mixing). Given a stationary time series  $(\mathbf{Y}_i)_{i \in \mathbb{Z}}$ , denote by  $\mathcal{F}_j^k$  the  $\sigma$ -field generated by  $(\mathbf{Y}_i)_{j \leq i \leq k}$ ,  $j, k \in \mathbb{Z} \cup \{-\infty, +\infty\}$ , and recall that the strong mixing coefficients corresponding to the stationary sequence  $(\mathbf{Y}_i)_{i \in \mathbb{Z}}$  are then defined by

$$\alpha_r^{\mathbf{Y}} = \sup_{A \in \mathcal{F}_{-\infty}^0, B \in \mathcal{F}_r^{+\infty}} |\mathbb{P}(A \cap B) - \mathbb{P}(A)\mathbb{P}(B)|, \quad r \in \mathbb{N}, r > 0,$$

with  $\alpha_0^{\mathbf{Y}} = 1/2$  and that the sequence  $(\mathbf{Y}_i)_{i \in \mathbb{Z}}$  is said to be strongly mixing if  $\alpha_r^{\mathbf{Y}} \rightarrow 0$  as  $r \rightarrow \infty$ .

*Remark 2.3.16.* Note that other conditions that entail [Condition 2.3.12](#) include, among others,  $\beta$ -mixing (implied by strongly mixing) (see, e.g., [Kosorok, 2008](#), [Chapter 11.6](#)) and  $\eta$ -dependence (see, e.g., [Doukhan et al., 2009](#), [Theorem 1](#)).

In another further step, [Bücher and Kojadinovic \(2016\)](#) extended the results in [Bücher and Volgushev \(2013\)](#) to the two sided empirical copula process  $\mathbb{C}_n$  in [\(2.37\)](#), and found that it suffices to impose the following condition.

**Condition 2.3.17** (Weak convergence of  $\mathbb{B}_n(0, \cdot, \cdot)$ ). The sequential empirical process  $\mathbb{B}_n(0, \cdot, \cdot)$  in [\(2.42\)](#) converges weakly in  $\ell^\infty([0, 1]^{d+1})$  to a tight centered Gaussian process  $\mathbb{Z}_C$  concentrated on

$$\{f \in \mathcal{C}([0, 1]^{d+1}) : f(s, \mathbf{u}) = 0 \text{ if one of the components of } (s, \mathbf{u}) \text{ is } 0, \text{ and} \\ f(s, 1, \dots, 1) = 0 \text{ for all } s \in (0, 1]\}.$$

Under Condition 2.3.17, it immediately follows from the continuous mapping theorem that  $\mathbb{B}_n \rightsquigarrow \mathbb{B}_C$  in  $\ell^\infty(\Delta \times [0, 1]^d)$ , where

$$\mathbb{B}_C(s, t, \mathbf{u}) = \mathbb{Z}_C(t, \mathbf{u}) - \mathbb{Z}_C(s, \mathbf{u}), \quad (s, t, \mathbf{u}) \in \Delta \times [0, 1]^d. \quad (2.43)$$

**Theorem 2.3.18** (Asymptotics of  $\mathbb{C}_n$ ). [*Bücher and Kojadinovic (2016)*] Suppose that the sequence  $(\mathbf{X}_i)_{i \in \mathbb{Z}}$  is a stationary time series that satisfies Condition 2.2.6 and 2.3.17 and that Condition 2.3.8 is satisfied. Then we have that

$$\sup_{(s, t, \mathbf{u}) \in \Delta \times [0, 1]^d} \left| \mathbb{C}_n(s, t, \mathbf{u}) - \tilde{\mathbb{C}}_n(s, t, \mathbf{u}) \right| = o_{\mathbb{P}}(1),$$

where

$$\tilde{\mathbb{C}}_n(s, t, \mathbf{u}) = \mathbb{B}_n(s, t, \mathbf{u}) - \sum_{j=1}^d \dot{C}_j(\mathbf{u}) \mathbb{B}_n(s, t, \mathbf{u}^{(j)}), \quad (s, t, \mathbf{u}) \in \Delta \times [0, 1]^d, \quad (2.44)$$

and  $\mathbb{B}_n$  is defined in (2.42). Consequently,  $\mathbb{C}_n \rightsquigarrow \mathbb{C}_C$  in  $\ell^\infty(\Delta \times [0, 1]^d)$ , where

$$\mathbb{C}_C(s, t, \mathbf{u}) = \mathbb{B}_C(s, t, \mathbf{u}) - \sum_{j=1}^d \dot{C}_j(\mathbf{u}) \mathbb{B}_C(s, t, \mathbf{u}^{(j)}), \quad (s, t, \mathbf{u}) \in \Delta \times [0, 1]^d, \quad (2.45)$$

and  $\mathbb{B}_C$  is defined in (2.43).

According to Theorem 2.1 in *Bücher (2015)*, Condition 2.3.17 is satisfied when  $(\mathbf{X}_i)_{i \in \mathbb{Z}}$  is strongly mixing. The next example shows how the asymptotics of  $\mathbb{C}_n$  could be used in *offline change-point detection*.

**Example 2.3.19** (Offline change-point detection). [*Bücher et al. (2014)*] In the time series setting, the empirical copula process can be applied to the offline change-point detection. Specifically, the goal is to test

$$H_0 : \mathcal{X}_{1:n} \text{ is a stretch from a stationary time series (of continuous random vectors)} \quad (2.46)$$

against

$$H_1 : \exists \text{ distinct } C_1, C_2 \text{ and } k^* \in \{1, \dots, n-1\} \text{ such that}$$

$$\mathbf{X}_1, \dots, \mathbf{X}_{k^*} \text{ have copula } C_1 \text{ and } \mathbf{X}_{k^*+1}, \dots, \mathbf{X}_n \text{ have copula } C_2.$$

For the purpose of the test above, *Bücher et al. (2014)* introduced the difference process  $\mathbb{D}_n$  defined by

$$\mathbb{D}_n(s, \mathbf{u}) = \sqrt{n} \lambda_n(0, s) \lambda_n(s, 1) \{C_{1:[ns]}(\mathbf{u}) - C_{[ns]+1:n}(\mathbf{u})\}, \quad (s, \mathbf{u}) \in [0, 1]^{d+1}, \quad (2.47)$$

where  $C_{k:l}$  is defined in (2.35), and proposed the test-statistic  $S_n$  defined by

$$S_n = \sup_{s \in [0,1]} \int_{[0,1]^d} \{\mathbb{D}_n(s, \mathbf{u})\}^2 dC_{1:n}(\mathbf{u}). \quad (2.48)$$

Under  $H_0$  in (2.46),  $S_n$  can be written as

$$S_n = \sup_{s \in [0,1]} \int_{[0,1]^d} \{\lambda_n(s, 1) \mathbb{C}_n(0, s, \mathbf{u}) - \lambda_n(0, s) \mathbb{C}_n(s, 1, \mathbf{u})\}^2 dC_{1:n}(\mathbf{u}), \quad (2.49)$$

where  $\mathbb{C}_n$  is defined in (2.37), and  $S_n \rightsquigarrow S$ , where

$$S = \sup_{s \in [0,1]} \int_{[0,1]^d} \{\mathbb{C}_C(0, s, \mathbf{u}) - s \mathbb{C}_C(0, 1, \mathbf{u})\}^2 dC(\mathbf{u}), \quad (2.50)$$

where  $\mathbb{C}_C$  is defined in (2.45).

### 2.3.3 The empirical Bernstein copula process and the empirical beta copula process

Following Segers et al. (2017), the following two smooth empirical copula processes can be considered:

$$\mathbb{C}_{n,\mathbf{m}}^{\text{Bern}}(\mathbf{u}) = \sqrt{n} \{C_{1:n,\mathbf{m}}^{\text{Bern}}(\mathbf{u}) - C(\mathbf{u})\}, \quad \mathbf{u} \in [0, 1]^d, \quad (2.51)$$

$$\mathbb{C}_n^{\text{beta}}(\mathbf{u}) = \sqrt{n} \{C_{1:n}^{\text{beta}}(\mathbf{u}) - C(\mathbf{u})\}, \quad \mathbf{u} \in [0, 1]^d, \quad (2.52)$$

where  $C_{1:n,\mathbf{m}}^{\text{Bern}}$  and  $C_{1:n}^{\text{beta}}$  are defined in (2.27) and (2.28), respectively.

**Theorem 2.3.20** (Asymptotics of  $\mathbb{C}_{n,\mathbf{m}}^{\text{Bern}}$  and  $\mathbb{C}_n^{\text{beta}}$ ). *[Segers et al. (2017)] Suppose that Condition 2.3.8 is satisfied and that  $\mathbb{C}_n(0, 1, \cdot) \rightsquigarrow \mathbb{C}_C(0, 1, \cdot)$  in  $\ell^\infty([0, 1]^d)$ , where  $\mathbb{C}_C(0, 1, \cdot)$  is the tight centered Gaussian process defined in (2.45). Let  $\mathbf{m} = \mathbf{m}(n)$  be multi-indices such that  $m_* = \min(m_1, \dots, m_d) \rightarrow \infty$  as  $n \rightarrow \infty$ . Then we have that*

$$\sup_{\mathbf{u} \in [0,1]^d} |\mathbb{C}_{n,\mathbf{m}}^{\text{Bern}}(\mathbf{u}) - \mathbb{C}_n(0, 1, \mathbf{u})| = o_{\mathbb{P}}(1),$$

where  $\mathbb{C}_n(0, 1, \cdot)$  is defined in (2.37). Consequently,  $\mathbb{C}_{n,\mathbf{m}}^{\text{Bern}} \rightsquigarrow \mathbb{C}_C(0, 1, \cdot)$  in  $\ell^\infty([0, 1]^d)$ . Furthermore, since  $\mathbb{C}_n^{\text{beta}}$  is a special case of  $\mathbb{C}_{n,\mathbf{m}}^{\text{Bern}}$  when  $\mathbf{m} = (n, \dots, n)$ , we have that

$$\sup_{\mathbf{u} \in [0,1]^d} |\mathbb{C}_n^{\text{beta}}(\mathbf{u}) - \mathbb{C}_n(0, 1, \mathbf{u})| = o_{\mathbb{P}}(1),$$

and that  $\mathbb{C}_n^{\text{beta}} \rightsquigarrow \mathbb{C}_C(0, 1, \cdot)$  in  $\ell^\infty([0, 1]^d)$ .

## 2.4 Bootstrapping empirical copula processes

As mentioned in the previous chapter, many test statistics (see e.g. Example 2.3.19) can be written as functionals of one of the considered (smooth) empirical copula processes. However, the limiting distribution of the latter is often unknown. For instance, according to Theorem 2.3.18, the distribution of  $\mathbb{C}_C$  in (2.45) depends on the unknown copula  $C$ , its first-order partial derivatives and the serial dependence properties of the unobservable sequence  $(\mathbf{U}_i)_{i \in \mathbb{Z}}$  in (2.11). Therefore, it is necessary to rely on resampling techniques to obtain replicates of the (smooth) empirical copula process. In this section, we present various bootstrapped (smooth) empirical copula processes and some theoretical results related to their asymptotic validity.

*Remark 2.4.1.* It is worth noticing that there exists test statistics written as functionals of empirical processes that have known limiting distributions under the null. For instance, if  $d = 1$  and  $\mathcal{X}_{1:n}$  is i.i.d, then the Kolmogorov-Smirnov statistic defined by  $\sqrt{n} \sup_{x \in \mathbb{R}} |F_{1:n}(x) - F(x)|$  has a known limiting distribution; see Feller (1948) for the d.f. of its limiting distribution. In this case, there is no need to rely on resampling techniques.

### 2.4.1 Mathematical preliminaries

In this section, let us recall the mode of convergence classically used to state asymptotic validity results of resampling techniques when dealing with empirical processes; see, e.g., van der Vaart and Wellner (2000, Chapter 2.9) or Kosorok (2008, Section 2.2.3). Let

$$BL_1 = \{h : \ell^\infty(T) \rightarrow [-1, 1] \text{ such that,}$$

$$\text{for all } x, y \in \ell^\infty(T), |h(x) - h(y)| \leq \sup_{\mathbf{u} \in T} |x(\mathbf{u}) - y(\mathbf{u})|\}.$$

For any stochastic process  $\mathbb{X}_n = \mathbb{X}_n(\mathcal{X}_{1:n})$  on  $\ell^\infty(T)$  constructed from  $\mathcal{X}_{1:n}$ , let  $\mathbb{X}_n^{[1]} = \mathbb{X}_n^{[1]}(\mathcal{X}_{1:n}, \mathcal{W}_n^{[1]})$ ,  $\mathbb{X}_n^{[2]} = \mathbb{X}_n^{[2]}(\mathcal{X}_{1:n}, \mathcal{W}_n^{[2]})$ ,  $\dots$  be a sequence of its bootstrapped replicates, where  $\mathcal{W}_n^{[1]}, \mathcal{W}_n^{[2]}, \dots$  are identically distributed random vectors that can be interpreted as *bootstrap weights*. It is also assumed that the sample paths of  $\mathbb{X}_n^{[1]}, \mathbb{X}_n^{[2]}, \dots$ , are in  $\ell^\infty(T)$ . For the purpose of showing asymptotic validity, we need to rely on the following definition.

**Definition 2.4.2** (“Weak convergence conditional on data in probability”). With the previous notation, we say that  $\mathbb{X}_n^{[1]}$  converges weakly to  $\mathbb{X}$ , and write  $\mathbb{X}_n^{[1]} \xrightarrow[\mathcal{W}]{\mathbb{P}} \mathbb{X}$  in  $\ell^\infty(T)$ , conditionally on  $\mathcal{X}_{1:n}$  in probability, if and only if

- $\sup_{h \in BL_1} |\mathbb{E}\{h(\mathbb{X}_n^{[1]})^* \mid \mathcal{X}_{1:n}\} - \mathbb{E}\{h(\mathbb{X})\}| \rightarrow 0$  in outer probability,
- $\mathbb{E}\{h(\mathbb{X}_n^{[1]})^* \mid \mathcal{X}_{1:n}\} - \mathbb{E}\{h(\mathbb{X}_n^{[1]})_* \mid \mathcal{X}_{1:n}\} \xrightarrow{P} 0$  for all  $h \in BL_1$ ,

where  $h(\mathbb{X}_n)^*$  and  $h(\mathbb{X}_n)_*$  denote the minimal measurable majorant and maximal measurable minorant with respect to  $(\mathcal{X}_{1:n}, \mathcal{W}_n^{[1]})$ .

**Lemma 2.4.3** (Lemma 3.1 of [Bücher and Kojadinovic \(2019\)](#)). *With the previous definitions, the following three assertions are equivalent:*

(a)  $\mathbb{X}_n^{[1]} \xrightarrow[\mathcal{W}]{\mathbb{P}} \mathbb{X}$  in  $\ell^\infty(T)$ .

(b)

$$(\mathbb{X}_n, \mathbb{X}_n^{[1]}, \mathbb{X}_n^{[2]}) \rightsquigarrow (\mathbb{X}, \mathbb{X}^{[1]}, \mathbb{X}^{[2]}) \quad \text{in } \{\ell^\infty(T)\}^3,$$

where  $\mathbb{X}, \mathbb{X}^{[1]}, \mathbb{X}^{[2]}$  are i.i.d.

(c) For any  $B \geq 2$ ,

$$(\mathbb{X}_n, \mathbb{X}_n^{[1]}, \dots, \mathbb{X}_n^{[B]}) \rightsquigarrow (\mathbb{X}, \mathbb{X}^{[1]}, \dots, \mathbb{X}^{[B]}) \quad \text{in } \{\ell^\infty(T)\}^{B+1},$$

where  $\mathbb{X}, \mathbb{X}^{[1]}, \dots, \mathbb{X}^{[B]}$  are i.i.d.

## 2.4.2 Bootstrapping the empirical copula processes

In this section, we first present the *straightforward bootstrap* and the *smooth bootstrap* proposed in [Kiriliouk et al. \(2021\)](#). Then, we present the *sequential dependent multiplier bootstrap* proposed in [Bücher and Kojadinovic \(2016\)](#).

### 2.4.2.1 Straightforward bootstraps in the i.i.d. case

In this section, we follow [Kiriliouk et al. \(2021\)](#) and present the straightforward bootstraps for  $\mathbb{C}_n(0, 1, \cdot)$  in (2.37) and  $\hat{\mathbb{C}}_n(0, 1, \cdot)$  in (2.38). First let  $\mathcal{W}_n^{[1]}, \mathcal{W}_n^{[2]}, \dots$ , be i.i.d. random vectors, which are independent of  $\mathbf{X}_{1:n}$ , and for  $b \in \mathbb{N}$ ,  $\mathcal{W}_n^{[b]} = (W_{n1}^{[b]}, \dots, W_{nn}^{[b]})$  is a multinomial random vector with probabilities  $(1/n, \dots, 1/n)$ . We first define

$$F_{1:n}^{[b]}(\mathbf{x}) = \frac{1}{n} \sum_{i=1}^n W_{ni}^{[b]} \mathbf{1}(\mathbf{X}_i \leq \mathbf{x}), \quad (2.53)$$

$$F_{1:n,j}^{[b]}(x_j) = \frac{1}{n} \sum_{i=1}^n W_{ni}^{[b]} \mathbf{1}(X_{ij} \leq x_j), \quad j \in \{1, \dots, d\}. \quad (2.54)$$

Then, by analogy with (2.17), we are able to define the bootstrapped (scaled) ranks

$$\hat{\mathbf{U}}_i^{1:n,[b]} = \mathbf{R}_i^{1:n,[b]} / n = \left( F_{n,1}^{1:n,[b]}(X_{i1}), \dots, F_{n,d}^{1:n,[b]}(X_{id}) \right), \quad i \in \{1, \dots, n\}. \quad (2.55)$$

Finally, the bootstrapped version of  $C_{1:n}$  in (2.20) and  $\hat{C}_{1:n}$  in (2.22) can be defined by

$$C_{1:n}^{[b]}(\mathbf{u}) = \frac{1}{n} \sum_{i=1}^n \mathbf{1}(\hat{U}_i^{1:n,[b]} \leq \mathbf{u}), \quad \mathbf{u} \in [0, 1]^d, \quad (2.56)$$

$$\hat{C}_{1:n}^{[b]}(\mathbf{u}) = F_{1:n}^{[b]} \left( F_{1:n,1}^{-1,[b]}(u_1), \dots, F_{1:n,d}^{-1,[b]}(u_d) \right), \quad \mathbf{u} \in [0, 1]^d, \quad (2.57)$$

where  $F_{1:n,j}^{-1,[b]}$  is the generalized inverse of  $F_{1:n,j}^{[b]}$  in (2.54). Then, for any  $b \in \mathbb{N}$ , the bootstrapped empirical copula processes are thus given by

$$\mathbb{C}_n^{[b]}(\mathbf{u}) = \sqrt{n} (C_{1:n}^{[b]}(\mathbf{u}) - C_{1:n}(\mathbf{u})), \quad \mathbf{u} \in [0, 1]^d, \quad (2.58)$$

$$\hat{\mathbb{C}}_n^{[b]}(\mathbf{u}) = \sqrt{n} (\hat{C}_{1:n}^{[b]}(\mathbf{u}) - \hat{C}_{1:n}(\mathbf{u})), \quad \mathbf{u} \in [0, 1]^d. \quad (2.59)$$

**Theorem 2.4.4** (Kiriliouk et al. (2021)). *Assume that the random vectors in  $\mathcal{X}_{1:n}$  are i.i.d, and that Condition 2.3.8 holds. Then, for any  $b \in \mathbb{N}$ ,*

$$\mathbb{C}_n^{[b]} \xrightarrow[\mathcal{W}]{\mathbb{P}} \mathbb{G}_C \quad \text{and} \quad \hat{\mathbb{C}}_n^{[b]} \xrightarrow[\mathcal{W}]{\mathbb{P}} \mathbb{G}_C$$

in  $\ell^\infty([0, 1]^d)$ , where  $\mathbb{G}_C$  is defined in (2.39).

### 2.4.2.2 Smooth bootstraps in the i.i.d. case

In this section, we follow Kiriliouk et al. (2021) and present the smooth bootstraps for  $\mathbb{C}_n(0, 1, \cdot)$  in (2.37),  $\hat{\mathbb{C}}_n(0, 1, \cdot)$  in (2.38) and  $\mathbb{C}_n^{\text{beta}}$  in (2.52). As can be seen from (2.54) and (2.55), the bootstrapped sample  $\hat{U}_i^{1:n,[b]}$  in (2.55) can be understood as sampling from  $C_{1:n}$  in (2.20). This idea can be applied to  $C_{1:n}^{\text{beta}}$  once we know how to sample from it. It follows from (2.29) that  $C_{1:n}^{\text{beta}}$  is a mixture of  $n$   $d$ -dimensional distributions having beta margins and whose copula is the independence copula. Then, to generate one random variate from  $C_{1:n}^{\text{beta}}$ , it thus suffices to randomly select one of the  $n$  components of the mixture by drawing a uniform on  $\{1, \dots, n\}$  and then generate one random variate from the selected  $d$ -dimensional distribution.

**Algorithm 2.4.5** (Sampling from  $C_{1:n}^{\text{beta}}$ ). [Algorithm 3.2 of Kiriliouk et al. (2021)]

1. Generate  $I$  from the discrete uniform distribution on  $\{1, \dots, n\}$ .
2. Independently generate  $U_j^\#$  from  $\text{Beta}(R_{Ij}^{1:n}, n + 1 - R_{Ij}^{1:n})$  for  $j \in \{1, \dots, d\}$ .
3. Set  $\mathbf{U}^\# = (U_1^\#, \dots, U_d^\#)$ .

For  $b \in \mathbb{N}$ , let  $\mathcal{U}_{1:n}^{\#,[b]} = (\mathbf{U}_1^{\#,[b]}, \dots, \mathbf{U}_n^{\#,[b]})$  be a random sample of size  $n$  from  $C_{1:n}^{\text{beta}}$  obtained by applying Algorithm 2.4.5  $n$  times independently. Then for any  $b \in \mathbb{N}$ , let  $G_{1:n}^{\#,[b]}$  be the empirical d.f. computed from the  $\mathcal{U}_{1:n}^{\#,[b]}$  and, for any  $j \in \{1, \dots, d\}$ , let  $G_{1:n,j}^{\#,[b]}$  be the empirical d.f. computed from the  $j$ th component sample  $U_{1j}^{\#,[b]}, \dots, U_{nj}^{\#,[b]}$  of  $\mathcal{U}_{1:n}^{\#,[b]}$ . Then,  $R_{ij}^{1:n,\#,[b]} = nG_{1:n,j}^{\#,[b]}(U_{ij}^\#)$  is the

rank of  $U_{ij}^{\#, [b]}$  among  $U_{1j}^{\#, [b]}, \dots, U_{nj}^{\#, [b]}$ . The bootstrapped version of  $C_{1:n}$  for  $\mathcal{U}_{1:n}^{\#, [b]}$  is thus given by

$$C_{1:n}^{\#, [b]}(\mathbf{u}) = \frac{1}{n} \sum_{i=1}^n \prod_{j=1}^d \mathbf{1} \left( \frac{R_{ij}^{1:n, \#, [b]}}{n} \leq u_j \right), \quad \mathbf{u} \in [0, 1]^d, \quad (2.60)$$

The bootstrapped version of  $\hat{C}_{1:n}$  for  $\mathcal{U}_{1:n}^{\#, [b]}$  is thus given by

$$\hat{C}_{1:n}^{\#, [b]}(\mathbf{u}) = G_{1:n}^{\#, [b]} \left( G_{1:n,1}^{-1, \#, [b]}(u_1), \dots, G_{1:n,d}^{-1, \#, [b]}(u_d) \right), \quad \mathbf{u} \in [0, 1]^d, \quad (2.61)$$

and the bootstrapped version of  $C_{1:n}^{\text{beta}}$  for  $\mathcal{U}_{1:n}^{\#, [b]}$  is

$$C_{1:n}^{\#, \text{beta}, [b]}(\mathbf{u}) = \frac{1}{n} \sum_{i=1}^n \prod_{j=1}^d \text{beta}_{n, R_{ij}^{1:n, \#, [b]}}(u_j), \quad \mathbf{u} \in [0, 1]^d. \quad (2.62)$$

Then, for any  $b \in \mathbb{N}$ , we are able to define the bootstrapped empirical copula processes:

$$\mathbb{C}_n^{\#, [b]}(\mathbf{u}) = \sqrt{n} (C_{1:n}^{\#, [b]}(\mathbf{u}) - C_{1:n}(\mathbf{u})), \quad \mathbf{u} \in [0, 1]^d, \quad (2.63)$$

$$\hat{\mathbb{C}}_n^{\#, [b]}(\mathbf{u}) = \sqrt{n} (\hat{C}_{1:n}^{\#, [b]}(\mathbf{u}) - \hat{C}_{1:n}(\mathbf{u})), \quad \mathbf{u} \in [0, 1]^d, \quad (2.64)$$

$$\mathbb{C}_n^{\#, \text{beta}, [b]}(\mathbf{u}) = \sqrt{n} (C_{1:n}^{\#, \text{beta}, [b]}(\mathbf{u}) - C_{1:n}^{\text{beta}}(\mathbf{u})), \quad \mathbf{u} \in [0, 1]^d. \quad (2.65)$$

**Theorem 2.4.6** (Kiriliouk et al. (2021)). *Assume that the random vectors in  $\mathcal{X}_{1:n}$  are i.i.d, and that Condition 2.3.8 holds. Then, for any  $b \in \mathbb{N}$ ,*

$$\mathbb{C}_n^{\#, [b]} \xrightarrow[\mathcal{W}]{\mathbb{P}} \mathbb{G}_C, \quad \hat{\mathbb{C}}_n^{\#, [b]} \xrightarrow[\mathcal{W}]{\mathbb{P}} \mathbb{G}_C \quad \text{and} \quad \mathbb{C}_n^{\#, \text{beta}, [b]} \xrightarrow[\mathcal{W}]{\mathbb{P}} \mathbb{G}_C$$

in  $\ell^\infty([0, 1]^d)$ , where  $\mathbb{G}_C$  defined in (2.39).

### 2.4.2.3 Sequential dependent multiplier bootstrap for $\mathbb{C}_n$ in the time series case

In this section, we follow Bücher and Kojadinovic (2016) and present the sequential multiplier bootstraps for  $\mathbb{C}_n$  in (2.37). The main idea can be summarized as follows: according to Theorem 2.3.18, in order to bootstrap  $\mathbb{C}_n$  in (2.37) in an asymptotically valid way, we simply need to bootstrap the process  $\tilde{\mathbb{C}}_n$  in (2.44). The latter could be done by bootstrapping  $\mathbb{B}_n$  in (2.42) and estimating the first-order partial derivatives  $\dot{C}_j$ ,  $j \in \{1, \dots, d\}$ , of  $C$ .

In the independent and non-sequential setting, to resample  $\mathbb{B}_n(0, 1, \cdot)$ , Scaillet (2005) and Rémillard and Scaillet (2009) used a *multiplier bootstrap* in the spirit of van der Vaart and Wellner (2000, Chapter 2.9). Specifically, the idea is to replace the random weights  $(W_{n1}, \dots, W_{nn})$  in Section 2.4.2.1 by  $(\xi_{1,n}, \dots, \xi_{n,n})$  which comes from an *i.i.d. multiplier sequence* given by the definition below.

**Definition 2.4.7** (i.i.d. multiplier sequence). We say that a sequence of random variables  $(\xi_{i,n})_{i \in \mathbb{Z}}$  is an i.i.d. multiplier sequence if:

(M0)  $(\xi_{i,n})_{i \in \mathbb{Z}}$  is i.i.d, independent of  $\mathcal{X}_{1:n}$ , with distribution not changing with  $n$ , having mean 0, variance 1, and being such that  $\int_0^\infty \{\mathbb{P}(|\xi_{0,n}| > x)\}^{1/2} dx < \infty$ .

Although the i.i.d. multiplier bootstrap frequently appears to lead to inference procedures with a good finite-sample performance, it cannot be applied to the time series case due to its lack of ability to capture the serial dependence in the data. Therefore, a *dependent multiplier sequence* is needed and its intuition can be summarized as follows. For the resampling in the time series case, [Künsch \(1989\)](#) proposed a method called the *block bootstrap* by drawing blocks of length  $l_n$  (rather than single observations) from the sample each time. Assume that  $n = kl_n$  (otherwise the last block is truncated) and for  $j \in \{1, \dots, k\}$ , let  $S_j$  be uniformly distributed on  $\{0, \dots, n - l_n\}$ . The block bootstrap sample is then given by

$$\mathbf{X}_{S_1+1}, \dots, \mathbf{X}_{S_1+l_n}, \mathbf{X}_{S_2+1}, \dots, \mathbf{X}_{S_2+l_n}, \dots, \mathbf{X}_{S_k+1}, \dots, \mathbf{X}_{S_k+l_n}.$$

Then, according to Section 3.3 in [Bühlmann \(1993\)](#), the blockwise bootstrapped empirical d.f. at  $\mathbf{x} \in \mathbb{R}^d$  is given by

$$F_{1:n}^{block}(\mathbf{x}) = \frac{1}{n} \sum_{i=1}^k \sum_{j=S_i+1}^{S_i+l_n} \mathbf{1}(\mathbf{X}_j \leq \mathbf{x}),$$

which can be equivalently expressed as

$$F_{1:n}^{block}(\mathbf{x}) = \frac{1}{n} \sum_{j=1}^n \zeta_{j,n} \mathbf{1}(\mathbf{X}_j \leq \mathbf{x}), \quad (2.66)$$

where  $\zeta_{j,n} = \sum_{i=1}^k \mathbf{1}(j - l_n \leq S_i \leq j - 1)$ . Interestingly, for  $j, h$  not at the border and  $n$  sufficiently large,  $\text{Cov}(\zeta_j, \zeta_h) \approx \max(1 - |j - h|/l_n, 0)$ . The *dependent multiplier sequence* can be regarded as an extension of the random weights  $\zeta_{j,n}$  in (2.66); see also [Chen and Fan \(1999\)](#), [Bücher and Ruppert \(2013\)](#), [Bücher et al. \(2014\)](#) and [Bücher and Kojadinovic \(2016\)](#).

**Definition 2.4.8** (Dependent multiplier sequence). We say that a sequence of random variables  $(\xi_{i,n})_{i \in \mathbb{Z}}$  is a dependent multiplier sequence if:

(M1) The sequence of random variables  $(\xi_{i,n})_{i \in \mathbb{Z}}$  is stationary with  $\mathbb{E}(\xi_{0,n}) = 0$ ,  $\mathbb{E}(\xi_{0,n}^2) = 1$  and  $\sup_{n \geq 1} \mathbb{E}(|\xi_{0,n}|^\gamma) < \infty$  for all  $\gamma \geq 1$ , and is independent of the available sample  $\mathcal{X}_{1:n}$ .

(M2) There exists a sequence  $l_n \rightarrow \infty$  of strictly positive constants such that  $l_n = o(n)$  and the sequence  $(\xi_{i,n})_{i \in \mathbb{Z}}$  is  $l_n$ -dependent, i.e.,  $\xi_{i,n}$  is independent of  $\xi_{i+h,n}$  for all  $h > l_n$  and  $i \in \mathbb{N}$ .

(M3) There exists a function  $\varphi : \mathbb{R} \rightarrow [0, 1]$ , symmetric around 0, continuous at 0, satisfying  $\varphi(0) = 1$  and  $\varphi(x) = 0$  for all  $|x| > 1$  such that  $\mathbb{E}(\xi_{0,n} \xi_{h,n}) = \varphi(h/l_n)$  for all  $h \in \mathbb{Z}$ .

Then, let  $(\xi_{i,n}^{[1]})_{i \in \mathbb{Z}}$ ,  $(\xi_{i,n}^{[2]})_{i \in \mathbb{Z}}, \dots$ , be independent copies of the same multiplier sequence. Two different multiplier bootstrap replicates of the process  $\mathbb{B}_n$  in (2.42) were proposed in [Bücher](#)



and [Kojadinovic \(2016\)](#) and [Bücher et al. \(2014\)](#), respectively. For any  $b \in \mathbb{N}$ ,  $(s, t) \in \Delta$  and  $\mathbf{u} \in [0, 1]^d$ , they are defined by

$$\hat{\mathbb{B}}_n^{[b]}(s, t, \mathbf{u}) = \frac{1}{\sqrt{n}} \sum_{i=[ns]+1}^{[nt]} \xi_{i,n}^{[b]} \left\{ \mathbf{1}(\hat{U}_i^{1:n} \leq \mathbf{u}) - C_{1:n}(\mathbf{u}) \right\} \quad (2.67)$$

and

$$\check{\mathbb{B}}_n^{[b]}(s, t, \mathbf{u}) = \frac{1}{\sqrt{n}} \sum_{i=[ns]+1}^{[nt]} \xi_{i,n}^{[b]} \left\{ \mathbf{1}(\hat{U}_i^{[ns]+1:[nt]} \leq \mathbf{u}) - C_{[ns]+1:[nt]}(\mathbf{u}) \right\}, \quad (2.68)$$

respectively, where  $C_{1:n}$  and  $C_{[ns]+1:[nt]}$  are generically defined in [\(2.35\)](#) and with the convention that  $\hat{\mathbb{B}}_n^{[b]}(s, t, \cdot) = \check{\mathbb{B}}_n^{[b]}(s, t, \cdot) = 0$  if  $[nt] - [ns] = 0$ .

In order to define multiplier bootstrap replicates of  $\tilde{\mathbb{C}}_n$  in [\(2.44\)](#), it is further necessary to estimate the unknown first-order partial derivatives  $\dot{C}_j$ ,  $j \in \{1, \dots, d\}$ , of  $C$ . In the rest of this section,  $\check{C}_{j,k:l}$  will denote an estimator of  $\dot{C}_j$  based on a stretch  $\mathbf{X}_{k:l} = (\mathbf{X}_k, \dots, \mathbf{X}_l)$  of observations,  $1 \leq k \leq l \leq n$ , with the convention that  $\check{C}_{j,k:l} = 0$  if  $k > l$ . Then, following [Bücher and Kojadinovic \(2016\)](#) and [Bücher et al. \(2014\)](#), we consider two types of multiplier bootstrap replicates of  $\mathbb{C}_n$  in [\(2.37\)](#). For any  $b \in \mathbb{N}$ ,  $(s, t) \in \Delta$  and  $\mathbf{u} \in [0, 1]^d$ , these are defined by

$$\hat{\mathbb{C}}_n^{[b]}(s, t, \mathbf{u}) = \hat{\mathbb{B}}_n^{[b]}(s, t, \mathbf{u}) - \sum_{j=1}^d \dot{C}_{j,1:n}(\mathbf{u}) \hat{\mathbb{B}}_n^{[b]}(s, t, \mathbf{u}^{(j)}) \quad (2.69)$$

and

$$\check{\mathbb{C}}_n^{[b]}(s, t, \mathbf{u}) = \check{\mathbb{B}}_n^{[b]}(s, t, \mathbf{u}) - \sum_{j=1}^d \dot{C}_{j,[ns]+1:[nt]}(\mathbf{u}) \check{\mathbb{B}}_n^{[b]}(s, t, \mathbf{u}^{(j)}), \quad (2.70)$$

respectively, where  $\hat{\mathbb{B}}_n^{[b]}$  (resp.  $\check{\mathbb{B}}_n^{[b]}$ ) is defined in [\(2.67\)](#) (resp. [\(2.68\)](#)). Clearly, both types of replicates coincide in a non-sequential setting as  $\hat{\mathbb{C}}_n^{[b]}(0, 1, \cdot) = \check{\mathbb{C}}_n^{[b]}(0, 1, \cdot)$ .

**Condition 2.4.9** (Bounded and weakly consistent partial derivative estimators). There exists a constant  $\zeta > 0$  such that, for any  $j \in \{1, \dots, d\}$  and  $n \in \mathbb{N}$ ,

$$\sup_{(s,t,\mathbf{u}) \in \Delta \times [0,1]^d} \left| \dot{C}_{j,[ns]+1:[nt]}(\mathbf{u}) \right| \leq \zeta.$$

Furthermore, for any  $\delta \in (0, 1)$ ,  $\varepsilon \in (0, 1/2)$  and  $j \in \{1, \dots, d\}$ ,

$$\sup_{\substack{(s,t) \in \Delta \\ t-s \geq \delta}} \sup_{\substack{\mathbf{u} \in [0,1]^d \\ u_j \in [\varepsilon, 1-\varepsilon]}} \left| \dot{C}_{j,[ns]+1:[nt]}(\mathbf{u}) - \dot{C}_j(\mathbf{u}) \right| = o_{\mathbb{P}}(1).$$

[Bücher and Kojadinovic \(2016\)](#) imposed the following condition on the observations and the underlying multiplier sequences.

**Condition 2.4.10** (Strong mixing and multiplier conditions). One of the following two conditions holds:

- (i) The random vectors in  $\mathcal{X}_{1:n}$  are i.i.d. and  $(\xi_{i,n}^{[1]})_{i \in \mathbb{Z}}, (\xi_{i,n}^{[2]})_{i \in \mathbb{Z}}, \dots$  are independent copies of a multiplier sequence satisfying (M0).
- (ii) The stretch  $\mathcal{X}_{1:n}$  is drawn from a stationary sequence  $(\mathbf{X}_i)_{i \in \mathbb{Z}}$  whose strong mixing coefficients satisfy  $\alpha_r^{\mathbf{X}} = O(r^{-a})$  for some  $a > 3 + 3d/2$  as  $r \rightarrow \infty$ . Furthermore,  $(\xi_{i,n}^{[1]})_{i \in \mathbb{Z}}, (\xi_{i,n}^{[2]})_{i \in \mathbb{Z}}, \dots$  are independent copies of a dependent multiplier sequence satisfying (M1)–(M3) with  $\ell_n = O(n^{1/2-\gamma})$  for some  $0 < \gamma < 1/2$ .

**Theorem 2.4.11** (Bücher and Kojadinovic (2016)). *Under Condition 2.3.8, 2.4.9 and 2.4.10,*

$$\begin{aligned} (\mathbb{C}_n, \hat{\mathbb{C}}_n^{[1]}, \hat{\mathbb{C}}_n^{[2]}) &\rightsquigarrow (\mathbb{C}_C, \mathbb{C}_C^{[1]}, \mathbb{C}_C^{[2]}), \\ (\mathbb{C}_n, \check{\mathbb{C}}_n^{[1]}, \check{\mathbb{C}}_n^{[2]}) &\rightsquigarrow (\mathbb{C}_C, \mathbb{C}_C^{[1]}, \mathbb{C}_C^{[2]}) \end{aligned}$$

in  $\{\ell^\infty(\Delta \times [0, 1]^d)\}^3$ , where  $\mathbb{C}_C^{[1]}$  and  $\mathbb{C}_C^{[2]}$  are independent copies of  $\mathbb{C}_C$  defined in (2.45).

The next example shows how the sequential dependent multiplier bootstrap for  $\mathbb{C}_n$  can be used to obtain replicates of the test statistic  $S_n$  under  $H_0$  in (2.46) and thus the approximate  $P$ -values.

**Example 2.4.12** (Offline change-point detection following Example 2.3.19). [Bücher et al. (2014)] In order to calculate the approximate p-values of the test statistic  $S_n$  in (2.49), Bücher et al. (2014) proposed the replicates of  $\mathbb{D}_n$  in (2.47) under  $H_0$  in (2.46). Specifically, for any  $b \in \mathbb{N}$  and  $(s, \mathbf{u}) \in [0, 1]^{d+1}$ , let

$$\begin{aligned} \hat{\mathbb{D}}_n^{[b]}(s, \mathbf{u}) &= \lambda_n(s, 1) \hat{\mathbb{C}}_n^{[b]}(0, s, \mathbf{u}) - \lambda_n(0, s) \hat{\mathbb{C}}_n^{[b]}(s, 1, \mathbf{u}), \\ \check{\mathbb{D}}_n^{[b]}(s, \mathbf{u}) &= \lambda_n(s, 1) \check{\mathbb{C}}_n^{[b]}(0, s, \mathbf{u}) - \lambda_n(0, s) \check{\mathbb{C}}_n^{[b]}(s, 1, \mathbf{u}), \end{aligned}$$

where  $\hat{\mathbb{C}}_n^{[b]}$  (resp.  $\check{\mathbb{C}}_n^{[b]}$ ) is defined in (2.69) (resp. (2.70)). Then, the replicates of  $S_n$  in (2.49) under  $H_0$  are therefore given by, for any  $b \in \{1, \dots, B\}$ ,

$$\begin{aligned} \hat{S}_n^{[b]} &= \sup_{s \in [0, 1]} \int_{[0, 1]^d} \{\hat{\mathbb{D}}_n^{[b]}(s, \mathbf{u})\}^2 dC_{1:n}(\mathbf{u}), \\ \check{S}_n^{[b]} &= \sup_{s \in [0, 1]} \int_{[0, 1]^d} \{\check{\mathbb{D}}_n^{[b]}(s, \mathbf{u})\}^2 dC_{1:n}(\mathbf{u}). \end{aligned}$$

Under  $H_0$  in (2.46) and the conditions of Theorem 2.4.11,

$$\begin{aligned} (S_n, \hat{S}_n^{[1]}, \hat{S}_n^{[2]}) &\rightsquigarrow (S, S^{[1]}, S^{[2]}), \\ (S_n, \check{S}_n^{[1]}, \check{S}_n^{[2]}) &\rightsquigarrow (S, S^{[1]}, S^{[2]}), \end{aligned}$$

where  $S$  is defined in (2.50) and  $S^{[1]}$  and  $S^{[2]}$  are independent copies of  $S$ . Then, the approximate p-values of the test statistic  $S_n$  can be computed via either

$$\frac{1}{B} \sum_{b=1}^B \mathbf{1}(\hat{S}_n^{[b]} \geq S_n) \quad \text{or} \quad \frac{1}{B} \sum_{b=1}^B \mathbf{1}(\check{S}_n^{[b]} \geq S_n).$$

# Chapter 3

## A class of smooth, possibly data-adaptive nonparametric copula estimators containing the empirical beta copula

This chapter is based on [Kojadinovic and Yi \(2022\)](#): A class of smooth, possibly data-adaptive nonparametric copula estimators containing the empirical beta copula. *Journal of Multivariate Analysis*, in press. <http://arxiv.org/abs/2106.10726>.

3.1	Introduction . . . . .	35
3.2	A broad class of smooth, possibly data-adaptive empirical copulas . . . . .	37
3.3	Properties and general form of the smooth estimators . . . . .	40
3.3.1	Univariate margins of the smooth estimators . . . . .	40
3.3.2	General form of the smooth estimators . . . . .	41
3.3.3	Conditions for being multivariate d.f.s . . . . .	43
3.4	Estimators based on smoothing distributions with scaled binomial, scaled beta-binomial or beta margins . . . . .	43
3.4.1	Scaled binomial and beta-binomial smoothing survival margins . . . . .	44
3.4.2	Beta smoothing survival margins . . . . .	46
3.4.3	On the survival copula of the smoothing distributions . . . . .	48
3.5	Suggested data-adaptive alternatives to the empirical beta copula . . . . .	49
3.5.1	Experiments for understanding the influence of the spread and the shape of the smoothing distributions . . . . .	49

3.5.2	Suggested data-adaptive smooth estimators and their finite-sample performance . . . . .	52
3.6	Asymptotics of related sequential empirical processes . . . . .	57
3.7	Concluding remarks . . . . .	60
3.8	Proofs . . . . .	61
3.8.1	Proofs of Lemma 3.2.1, Proposition 3.3.2, Proposition 3.3.7 and Proposition 3.3.10 . . . . .	61
3.8.2	Proofs of Propositions 3.4.1, 3.4.2 and 3.4.4 . . . . .	63
3.8.3	Proofs of Theorem 3.6.4 . . . . .	66

## Abstract

A broad class of smooth, possibly data-adaptive nonparametric copula estimators that contains empirical Bernstein copulas introduced by Sancetta and Satchell (and thus the empirical beta copula proposed by Segers, Sibuya and Tsukahara) is studied. Within this class, a subclass of estimators that depend on a scalar parameter determining the amount of marginal smoothing and a functional parameter controlling the shape of the smoothing region is specifically considered. Empirical investigations of the influence of these parameters suggest to focus on two particular data-adaptive smooth copula estimators that were found to be uniformly better than the empirical beta copula in all of the considered Monte Carlo experiments. Finally, with future applications to change-point detection in mind, conditions under which related sequential empirical copula processes converge weakly are provided.

## 3.1 Introduction

Let  $\mathcal{X}_{1:n} = (\mathbf{X}_1, \dots, \mathbf{X}_n)$  be a stretch of  $d$ -dimensional random vectors from a stationary time series  $(\mathbf{X}_i)_{i \in \mathbb{Z}}$ . The distribution function (d.f.)  $F$  of  $\mathbf{X}_1$  is assumed to have continuous univariate margins  $F_1, \dots, F_d$ . As a consequence of a well-known theorem of Sklar (1959), the multivariate d.f.  $F$  can be expressed as

$$F(\mathbf{x}) = C\{F_1(x_1), \dots, F_d(x_d)\}, \quad \mathbf{x} \in \mathbb{R}^d,$$

in terms of a unique copula  $C$ , that is, a unique  $d$ -dimensional d.f. with standard uniform margins.

To carry out statistical inference on the unknown copula  $C$  using the available observations  $\mathcal{X}_{1:n}$ , it is often necessary to have at hand nonparametric estimators of  $C$ . The best known

such estimator is called the empirical copula [Deheuvels \(1979\)](#). Under a rather weak condition (see [Condition 3.6.5](#) in [Section 3.6](#)), the latter is asymptotically equivalent to the empirical d.f. of the multivariate ranks obtained from  $\mathcal{X}_{1:n}$  scaled by  $1/n$  which was studied in [Rüschendorf \(1976\)](#). Two smooth versions that are genuine copulas when there are no ties in the components samples of  $\mathcal{X}_{1:n}$  are the empirical checkerboard copula (see, e.g., [Genest et al., 2017](#); [Genest and Nešlehová, 2007](#), and the references therein) and the empirical beta copula proposed in [Segers et al. \(2017\)](#). The latter was found to have better small-sample properties than the former and the classical empirical copula in the Monte Carlo experiments reported in [Segers et al. \(2017\)](#).

In this work, we investigate extensions of the construction that allowed [Segers et al. \(2017\)](#) to study the asymptotics of empirical Bernstein copulas introduced in [Sancetta and Satchell \(2004\)](#), and thus of the empirical beta copula. The initial motivation for this undertaking stems from an early attempt to obtain sequential versions of the asymptotic results of [Segers et al. \(2017\)](#) (with an application to change-point detection in mind) during which it appeared that alternative ways of smoothing could be considered. In particular, we allow the underlying smoothing distributions to depend on the data, leading to a rather broad class of smooth, possibly data-adaptive empirical copulas.

The paper is organized as follows. In [Section 3.2](#), we define a broad class of smooth, possibly data-adaptive nonparametric copula estimators that contains empirical Bernstein copulas, and thus the empirical beta copula. Conditions under which such smooth estimators have standard uniform univariate margins, are multivariate d.f.s or are genuine copulas are provided in [Section 3.3](#). In [Section 3.4](#), we focus on a subclass of empirical copulas that depend on a scalar parameter that determines the amount of marginal smoothing and a functional parameter that controls the shape of the smoothing region in  $[0, 1]^d$ . Using an implementation for the R statistical environment [R Core Team \(2022\)](#) (available on the web page of the first author), we investigate the influence of these parameters through Monte Carlo experiments in [Section 3.5](#) and, as a result, we suggest to focus on two specific smooth data-adaptive copula estimators that were found to be uniformly better than the empirical beta copula in all of the considered Monte Carlo experiments. Finally, in [Section 3.6](#), we study the weak convergence of the sequential empirical copula processes related to the general class of smooth estimators proposed in [Section 3.2](#) and, in particular, establish conditions under which they are asymptotically equivalent to the classical sequential empirical copula process initially studied in [Bücher and Kojadinovic \(2016\)](#).

### 3.2 A broad class of smooth, possibly data-adaptive empirical copulas

For any  $j \in \{1, \dots, d\}$ , let  $F_{1:n,j}$  be the empirical d.f. computed from the  $j$ th component sample  $X_{1j}, \dots, X_{nj}$  of the available observations  $\mathcal{X}_{1:n}$ . Then, let

$$R_{ij}^{1:n} = nF_{1:n,j}(X_{ij}) = \sum_{t=1}^n \mathbf{1}(X_{tj} \leq X_{ij})$$

be the (maximal) rank of  $X_{ij}$  among  $X_{1j}, \dots, X_{nj}$ . Furthermore, let  $\mathbf{R}_i^{1:n} = (R_{i1}^{1:n}, \dots, R_{id}^{1:n})$  and  $\hat{U}_i^{1:n} = \mathbf{R}_i^{1:n}/n$ ,  $i \in \{1, \dots, n\}$ , be the multivariate ranks and the multivariate scaled ranks, respectively, obtained from  $\mathcal{X}_{1:n}$ . Note that the  $d$ -dimensional random vectors  $\hat{U}_1^{1:n}, \dots, \hat{U}_n^{1:n}$  are sometimes referred to as pseudo-observations from  $C$  (see, e.g., Hofert et al., 2018, Chapter 4). Following Rüschendorf (1976), the empirical copula  $C_{1:n}$  of  $\mathcal{X}_{1:n}$  at  $\mathbf{u} = (u_1, \dots, u_d) \in [0, 1]^d$  can then be defined by

$$C_{1:n}(\mathbf{u}) = \frac{1}{n} \sum_{i=1}^n \prod_{j=1}^d \mathbf{1}(R_{ij}^{1:n}/n \leq u_j) = \frac{1}{n} \sum_{i=1}^n \mathbf{1}(\hat{U}_i^{1:n} \leq \mathbf{u}), \quad (3.1)$$

where inequalities between vectors are to be understood componentwise.

A smooth version of the empirical copula  $C_{1:n}$  was proposed in Segers et al. (2017) by replacing indicator functions in (3.1) by d.f.s of particular beta distributions. Specifically, the empirical beta copula of  $\mathcal{X}_{1:n}$  is defined by

$$C_{1:n}^{\text{beta}}(\mathbf{u}) = \frac{1}{n} \sum_{i=1}^n \prod_{j=1}^d \text{beta}_{n, R_{ij}^{1:n}}(u_j), \quad \mathbf{u} = (u_1, \dots, u_d) \in [0, 1]^d, \quad (3.2)$$

where, for any  $n \in \mathbb{N}$  and  $r \in \{1, \dots, n\}$ ,  $\text{beta}_{n,r}$  denotes the d.f. of the distribution  $\text{Beta}(r, n + 1 - r)$  (the beta distribution with shape parameters  $\alpha = r$  and  $\beta = n + 1 - r$ ). When there are no ties in the component samples of  $\mathcal{X}_{1:n}$ , Lemma 2.6 in Segers et al. (2017) states that the empirical beta copula is actually a particular case of the empirical Bernstein copula introduced in Sancetta and Satchell (2004) and further studied in Janssen et al. (2012). Proposition 2.8 in Segers et al. (2017) additionally shows that the supremum distance between the empirical beta copula  $C_{1:n}^{\text{beta}}$  and the classical empirical copula  $C_{1:n}$  is  $O(n^{-1/2}(\ln n)^{1/2})$ , thereby suggesting that the empirical beta copula is a smoothing of  $C_{1:n}$  at approximately bandwidth  $O(n^{-1/2})$ ; see also Corollary 3.7 in Segers et al. (2017).

To study the asymptotics of the empirical beta copula, Segers et al. (2017) cleverly used the fact that it could be written as a mixture involving the classical empirical copula. For any  $n \in \mathbb{N}$  and  $\mathbf{u} \in [0, 1]^d$ , let  $\boldsymbol{\mu}_{n,\mathbf{u}}$  be the law of the  $d$ -dimensional random vector  $(S_{n,1,u_1}/n, \dots, S_{n,d,u_d}/n)$ , where  $S_{n,1,u_1}, \dots, S_{n,d,u_d}$  are independent random variables and, for each  $j \in \{1, \dots, d\}$ ,  $S_{n,j,u_j}$  is  $\text{Binomial}(n, u_j)$ , that is,  $S_{n,j,u_j}$  follows a binomial distribution with parameters  $n$  and  $u_j$ . The

following lemma proven in Appendix 3.8.1 is instrumental for understanding the approach of Segers et al. (2017).

**Lemma 3.2.1.** For any  $\mathbf{u} \in [0, 1]^d$ ,  $n \in \mathbb{N}$  and  $r_1, \dots, r_d \in \{1, \dots, n\}$ ,

$$\prod_{j=1}^d \text{beta}_{n,r_j}(u_j) = \int_{[0,1]^d} \prod_{j=1}^d \mathbf{1}(r_j/n \leq w_j) d\mu_{n,\mathbf{u}}(\mathbf{w}).$$

Using Lemma 3.2.1, by linearity of the integral, one immediately obtains that, for any  $\mathbf{u} \in [0, 1]^d$ ,

$$C_{1:n}^{\text{beta}}(\mathbf{u}) = \frac{1}{n} \sum_{i=1}^n \int_{[0,1]^d} \prod_{j=1}^d \mathbf{1}(R_{ij}^{1:n}/n \leq w_j) d\mu_{n,\mathbf{u}}(\mathbf{w}) = \int_{[0,1]^d} C_{1:n}(\mathbf{w}) d\mu_{n,\mathbf{u}}(\mathbf{w}). \quad (3.3)$$

The main aim of this work is to study generalizations of the empirical beta copula based on alternative smoothing distributions, possibly depending on the data. As we continue, for any  $m \in \mathbb{N}$ ,  $\mathbf{x} \in (\mathbb{R}^d)^m$  and  $\mathbf{u} \in [0, 1]^d$ ,  $\nu_{\mathbf{u}}^{\mathbf{x}}$  will denote a law on  $[0, 1]^d$  such that, for any  $j \in \{1, \dots, d\}$ ,  $\int_{[0,1]^d} w_j d\nu_{\mathbf{u}}^{\mathbf{x}}(\mathbf{w}) = u_j$ ; it is meant to be a generalization of  $\mu_{n,\mathbf{u}}$  in (3.3) that possibly depends on the data set  $\mathbf{x}$ . Furthermore, let  $p \geq d$  be a fixed integer, let  $\mathbf{U}$  be a  $p$ -dimensional random vector whose components are independent and standard uniform, and consider the following assumption.

**Condition 3.2.2** (Construction of smoothing random vectors). For any  $m \in \mathbb{N}$ ,  $\mathbf{x} \in (\mathbb{R}^d)^m$  and  $\mathbf{u} \in [0, 1]^d$ , there exists a function  $\mathcal{W}_{\mathbf{u}}^{\mathbf{x}} : [0, 1]^p \rightarrow [0, 1]^d$  such that  $\mathbf{W}_{\mathbf{u}}^{\mathbf{x}} = \mathcal{W}_{\mathbf{u}}^{\mathbf{x}}(\mathbf{U})$  is a  $[0, 1]^d$ -valued random vector with law  $\nu_{\mathbf{u}}^{\mathbf{x}}$ .

To be able to define, for any  $n \in \mathbb{N}$ ,  $\mathcal{X}_{1:n} = (\mathbf{X}_1, \dots, \mathbf{X}_n)$  and, for any  $m \leq n$ , the random vectors  $\mathbf{W}_{\mathbf{u}}^{\mathbf{x}}$ ,  $\mathbf{x} \in (\mathbb{R}^d)^m$ ,  $\mathbf{u} \in [0, 1]^d$ , on the same probability space (the case  $m < n$  will be needed in Section 3.6), we assume that the underlying probability space  $(\Omega, \mathcal{A}, \mathbb{P})$  has a product structure, that is,  $\Omega = \Omega_0 \times \Omega_1$  with probability measure  $\mathbb{P} = \mathbb{P}_0 \otimes \mathbb{P}_1$ , where  $\mathbb{P}_i$  denotes the probability measure on  $\Omega_i$ , such that, for any  $\omega \in \Omega$ ,  $\mathcal{X}_{1:n}(\omega)$  only depends on the first coordinate of  $\omega$  and  $\mathbf{U}(\omega)$  only depends on the second coordinate of  $\omega$ , implying in particular that  $\mathcal{X}_{1:n}$  and  $\mathbf{U}$  are independent. In that case, it can be verified using Fubini's theorem that  $(\mathbf{x}, A) \mapsto \mathbb{P}_1\{\mathcal{W}_{\mathbf{u}}^{\mathbf{x}}(\mathbf{U}) \in A\} = \mathbb{P}_1\{\mathbf{W}_{\mathbf{u}}^{\mathbf{x}} \in A\} = \nu_{\mathbf{u}}^{\mathbf{x}}(A)$  defines a regular version of the conditional distribution  $\mathbb{P}(\mathbf{W}_{\mathbf{u}}^{\mathcal{X}_{1:n}} \in \cdot \mid \mathcal{X}_{1:n})$  of  $\mathbf{W}_{\mathbf{u}}^{\mathcal{X}_{1:n}}$  given  $\mathcal{X}_{1:n}$ . As a consequence, in the rest of the paper, for an arbitrary real-valued function  $h$ ,  $\mathbb{E}\{h(\mathbf{W}_{\mathbf{u}}^{\mathcal{X}_{1:n}}) \mid \mathcal{X}_{1:n}\}$  is to be understood as  $\int_{[0,1]^d} h(\mathbf{w}) d\nu_{\mathbf{u}}^{\mathcal{X}_{1:n}}(\mathbf{w})$ .

As we continue, for any  $m \in \mathbb{N}$ ,  $\mathbf{x} \in (\mathbb{R}^d)^m$  and  $\mathbf{u} \in [0, 1]^d$ , the  $d$  components of the random vector  $\mathbf{W}_{\mathbf{u}}^{\mathbf{x}}$  will be denoted by  $W_{1,u_1}^{\mathbf{x}}, \dots, W_{d,u_d}^{\mathbf{x}}$  to indicate that the  $j$ th component of  $\mathbf{W}_{\mathbf{u}}^{\mathbf{x}}$  depends on  $u_j$  but not on  $u_1, \dots, u_{j-1}, u_{j+1}, \dots, u_d$ . It is however important to keep in mind that the joint distribution of  $W_{1,u_1}^{\mathbf{x}}, \dots, W_{d,u_d}^{\mathbf{x}}$  may still depend on  $\mathbf{u}$ . The fact that, for any  $\mathbf{x} \in (\mathbb{R}^d)^m$  and  $\mathbf{u} \in [0, 1]^d$ ,  $\mathbf{W}_{\mathbf{u}}^{\mathbf{x}}$  has its support included in  $[0, 1]^d$  with  $\mathbb{E}(\mathbf{W}_{\mathbf{u}}^{\mathbf{x}}) = \mathbf{u}$  implies

that, for any  $j \in \{1, \dots, d\}$ ,  $\text{Var}(W_{j,u_j}^{\mathbf{x}}) = 0$  if  $u_j \in \{0, 1\}$ . Note that, more generally, for any  $\mathbf{x} \in (\mathbb{R}^d)^m$ ,  $\mathbf{u} \in [0, 1]^d$  and  $j \in \{1, \dots, d\}$ , one has that

$$\text{Var}(W_{j,u_j}^{\mathbf{x}}) = \mathbb{E} \left\{ (W_{j,u_j}^{\mathbf{x}})^2 \right\} - u_j^2 \leq \mathbb{E}(W_{j,u_j}^{\mathbf{x}}) - u_j^2 = u_j(1 - u_j). \quad (3.4)$$

By analogy with (3.3), we then define alternative smooth versions of the empirical copula  $C_{1:n}$  of the sample  $\mathcal{X}_{1:n} = (\mathbf{X}_1, \dots, \mathbf{X}_n)$  in (3.1) by

$$C_{1:n}^{\nu}(\mathbf{u}) = \int_{[0,1]^d} C_{1:n}(\mathbf{w}) d\nu_{\mathbf{u}}^{\mathcal{X}_{1:n}}(\mathbf{w}), \quad \mathbf{u} \in [0, 1]^d. \quad (3.5)$$

Roughly speaking, for any  $\mathbf{u} \in [0, 1]^d$ ,  $C_{1:n}^{\nu}(\mathbf{u})$  can be thought of as a “weighted average” of  $C_{1:n}(\mathbf{w})$  for  $\mathbf{w}$  “in a neighborhood of  $\mathbf{u}$ ” according to the smoothing distribution  $\nu_{\mathbf{u}}^{\mathcal{X}_{1:n}}$  (that may depend on the available observations  $\mathcal{X}_{1:n}$ ).

*Remark 3.2.3.* Let us comment on Condition 3.2.2. For  $n \in \mathbb{N}$ ,  $\mathbf{x} \in (\mathbb{R}^d)^n$  and  $\mathbf{u} \in [0, 1]^d$ , let  $H$  be the d.f. corresponding to  $\nu_{\mathbf{u}}^{\mathbf{x}}$ , let  $D$  be a copula of  $H$  and let  $H_1, \dots, H_d$  be the  $d$  univariate margins of  $H$ . Assume for instance first that  $D$  is an absolutely continuous copula. Then, using the inverse of the well-known transformation of Rosenblatt (1952), it is possible to obtain, from the first  $d$  components of  $\mathbf{U}$ , a  $d$ -dimensional random vector  $\mathbf{V}$  with d.f.  $D$  (see, e.g., Hofert et al., 2018, Section 2.7 for more details). The random vector  $\mathbf{W}_{\mathbf{u}}^{\mathbf{x}}$  can then be defined by  $(H_1^{-1}(V_1), \dots, H_d^{-1}(V_d))$ , where  $H_1^{-1}, \dots, H_d^{-1}$  are the quantile functions (generalized inverses) obtained from the  $d$  univariate margins of  $H$ . As another example, assume that  $D$  is the empirical beta copula obtained from the data set  $\mathbf{x} \in (\mathbb{R}^d)^n$  whose component samples contain no ties. Let  $r_{ij}$ ,  $i \in \{1, \dots, n\}$ ,  $j \in \{1, \dots, d\}$ , be the corresponding multivariate ranks. Then, according to (Kiriliouk et al., 2021, Section 3.2), a  $d$ -dimensional random vector  $\mathbf{V}$  with d.f.  $D$  can be obtained by computing  $I = \lfloor nU_{d+1} \rfloor + 1$  and  $V_j = \text{beta}_{n, r_{Ij}}^{-1}(U_j)$ ,  $j \in \{1, \dots, d\}$ , where, as in (3.2),  $\text{beta}_{n,q}$  denotes the d.f. of the distribution  $\text{Beta}(q, n + 1 - q)$ . Finally, the random vector  $\mathbf{W}_{\mathbf{u}}^{\mathbf{x}}$  can again be defined by  $(H_1^{-1}(V_1), \dots, H_d^{-1}(V_d))$ . More generally, as soon as there exists a method to generate random variates from the copula  $D$ , it is likely that we will be able to define the function  $\mathcal{W}_{\mathbf{u}}^{\mathbf{x}}$  transforming  $\mathbf{U}$  into the random vector  $\mathbf{W}_{\mathbf{u}}^{\mathbf{x}}$  with law  $\nu_{\mathbf{u}}^{\mathbf{x}}$ .

*Remark 3.2.4.* Let  $n, m_1, \dots, m_d \in \mathbb{N}$ . For any  $\mathbf{x} \in (\mathbb{R}^d)^n$  and  $\mathbf{u} \in [0, 1]^d$ , let  $\nu_{\mathbf{u}}^{\mathbf{x}}$  be the law of the random vector  $(S_{m_1,1,u_1}/m_1, \dots, S_{m_d,d,u_d}/m_d)$ , where  $S_{m_1,1,u_1}, \dots, S_{m_d,d,u_d}$  are independent random variables and, for each  $j \in \{1, \dots, d\}$ ,  $S_{m_j,j,u_j}$  is Binomial( $m_j, u_j$ ). Then, from Section 3 in Segers et al. (2017),  $C_{1:n}^{\nu}$  corresponds to the empirical Bernstein copula with polynomial degrees  $m_1, \dots, m_d$ . Since, from Lemma 2.6 in Segers et al. (2017), the empirical beta copula is the empirical Bernstein copula with  $m_1 = n, \dots, m_d = n$ , it is obviously a particular case of  $C_{1:n}^{\nu}$ . It is obtained when, for any  $\mathbf{x} \in (\mathbb{R}^d)^n$  and  $\mathbf{u} \in [0, 1]^d$ ,  $\nu_{\mathbf{u}}^{\mathbf{x}}$  is defined as the measure  $\mu_{n,\mathbf{u}}$  appearing in (3.3) which is the law of the random vector  $(S_{n,1,u_1}/n, \dots, S_{n,d,u_d}/n)$ .

As we shall see in Section 3.4, when defining some specific members of the above general class of smooth copulas, we will consider the possibly random smoothing distributions  $\nu_{\mathbf{u}}^{\mathcal{X}_{1:n}}$ ,  $\mathbf{u} \in [0, 1]^d$ ,



to be either all discrete or all continuous with the understanding that, because of (3.4), if  $u_j \in \{0, 1\}$ , the  $j$ th component of  $\mathbf{W}_u^{\mathbf{x}^{1:n}}$  will be degenerate (non-random).

For any  $n \in \mathbb{N}$ ,  $\mathbf{x} \in (\mathbb{R}^d)^n$ ,  $\mathbf{r} \in [0, n]^d$  and  $\mathbf{u} \in [0, 1]^d$ , let

$$\mathcal{K}_r^{\mathbf{x}}(\mathbf{u}) = \int_{[0,1]^d} \mathbf{1}(\mathbf{r}/n \leq \mathbf{w}) d\nu_{\mathbf{u}}^{\mathbf{x}}(\mathbf{w}) = \mathbb{E} \{ \mathbf{1}(\mathbf{r}/n \leq \mathbf{W}_u^{\mathbf{x}}) \} = \mathcal{G}_u^{\mathbf{x}}(\mathbf{r}/n), \quad (3.6)$$

where  $\mathcal{G}_u^{\mathbf{x}}(\mathbf{w}) = \mathbb{P}(\mathbf{W}_u^{\mathbf{x}} \geq \mathbf{w})$ ,  $\mathbf{w} \in [0, 1]^d$ . Note that this implies that, almost surely,  $\mathcal{G}_u^{\mathbf{x}^{1:n}}(\mathbf{w}) = \mathbb{P}(\mathbf{W}_u^{\mathbf{x}^{1:n}} \geq \mathbf{w} \mid \mathcal{X}_{1:n})$ ,  $\mathbf{w} \in [0, 1]^d$ . By linearity of the integral, with probability 1, we can then express  $C_{1:n}^{\nu}$  as

$$C_{1:n}^{\nu}(\mathbf{u}) = \frac{1}{n} \sum_{i=1}^n \mathcal{K}_{\mathbf{R}_i^{1:n}}^{\mathbf{x}^{1:n}}(\mathbf{u}), \quad \mathbf{u} \in [0, 1]^d. \quad (3.7)$$

### 3.3 Properties and general form of the smooth estimators

In this section, we provide conditions under which smooth empirical copulas of the form (3.5) have standard uniform univariate margins and then conditions under which they are multivariate d.f.s.

#### 3.3.1 Univariate margins of the smooth estimators

We start by investigating the univariate margins of the studied nonparametric copula estimators. As already hinted at in the introduction, the following simple condition plays an important role.

**Condition 3.3.1** (No ties). With probability 1, there are no ties in each of the component samples  $X_{1j}, \dots, X_{nj}$ ,  $j \in \{1, \dots, d\}$ , of  $\mathcal{X}_{1:n}$ .

As verified in Segers et al. (2017), under this condition, the empirical beta copula  $C_{1:n}^{\text{beta}}$  defined in (3.2) has standard uniform margins. In the rest of the paper, for any  $j \in \{1, \dots, d\}$  and any  $\mathbf{u} \in [0, 1]^d$ , let  $\mathbf{u}^{(j)}$  be the vector of  $[0, 1]^d$  defined by  $u_i^{(j)} = u_j$  if  $i = j$  and 1 otherwise. With this notation, the property of having standard uniform univariate margins can be simply written as  $C_{1:n}^{\text{beta}}(\mathbf{u}^{(j)}) = u_j$  for all  $j \in \{1, \dots, d\}$  and  $\mathbf{u} \in [0, 1]^d$ . For the smooth empirical copula  $C_{1:n}^{\nu}$  defined in (3.5), we have the following result proven in Appendix 3.8.1.

**Proposition 3.3.2** (Univariate margins of  $C_{1:n}^{\nu}$ ). *Under Condition 3.3.1, for any  $j \in \{1, \dots, d\}$  and  $\mathbf{u} \in [0, 1]^d$ , with probability 1,*

$$C_{1:n}^{\nu}(\mathbf{u}^{(j)}) = \mathbb{E} \left( \frac{\lfloor nW_{j,u_j}^{\mathbf{x}^{1:n}} \rfloor}{n} \mid \mathcal{X}_{1:n} \right).$$

Since  $\sup_{w \in [0,1]} |\lfloor nw \rfloor / n - w| \rightarrow 0$  as  $n \rightarrow \infty$  and  $\mathbb{E} \left( W_{j,u_j}^{\mathbf{x}^{1:n}} \mid \mathcal{X}_{1:n} \right) = u_j$  almost surely by construction, under Condition 3.3.1, a smooth empirical copula  $C_{1:n}^{\nu}$  will at least have standard

uniform margins asymptotically. Actually, it is easy to verify that, under Condition 3.3.1,  $C_{1:n}^\nu$  will have standard uniform margins if and only if the following condition is satisfied.

**Condition 3.3.3** (Condition for uniform margins). For any  $\mathbf{x} \in (\mathbb{R}^d)^n$ ,  $\mathbf{u} \in [0, 1]^d$  and  $j \in \{1, \dots, d\}$ ,  $W_{j,u_j}^{\mathbf{x}}$  takes its values in the set  $\{0, 1/n, \dots, (n-1)/n, 1\}$ .

The previous condition is for instance satisfied when, for any  $\mathbf{x} \in (\mathbb{R}^d)^n$ ,  $\mathbf{u} \in [0, 1]^d$  and  $j \in \{1, \dots, d\}$ ,  $W_{j,u_j}^{\mathbf{x}} = S_{n,j,u_j}/n$ , where  $S_{n,j,u_j}$  is Binomial( $n, u_j$ ). Notice that, when  $u_j \in (0, 1)$  and  $W_{j,u_j}^{\mathbf{x}_{1:n}}$  does not take all its values in the set  $\{0, 1/n, \dots, (n-1)/n, 1\}$ , one has with probability 1 that

$$C_{1:n}^\nu(\mathbf{u}^{(j)}) = \mathbb{E} \left( \frac{\lfloor nW_{j,u_j}^{\mathbf{x}_{1:n}} \rfloor}{n} \mid \mathcal{X}_{1:n} \right) < \mathbb{E} \left( W_{j,u_j}^{\mathbf{x}_{1:n}} \mid \mathcal{X}_{1:n} \right) = u_j$$

since  $\lfloor nw \rfloor/n < w$  for all  $w \in [0, 1] \setminus \{0, 1/n, \dots, (n-1)/n, 1\}$ . Hence, in that case, from a marginal perspective,  $C_{1:n}^\nu$  will systematically underestimate  $C$ .

One possible remedy to this situation is to carry out an asymptotically negligible correction consisting of using  $\mathcal{K}_{\mathbf{R}_i^{1:n} - \mathbf{a}}^{\mathbf{x}_{1:n}}$  instead of  $\mathcal{K}_{\mathbf{R}_i^{1:n}}^{\mathbf{x}_{1:n}}$  in (3.7), where  $\mathbf{a} = (a, \dots, a) \in (0, 1)^d$ . Indeed, mimicking the proof of Proposition 3.3.2, one obtains that, under Condition 3.3.1, for any  $j \in \{1, \dots, d\}$  and  $\mathbf{u} \in [0, 1]^d$ ,  $C_{1:n}^\nu(\mathbf{u}^{(j)}) = \mathbb{E}(\lfloor nW_{j,u_j}^{\mathbf{x}_{1:n}} + a \rfloor/n \mid \mathcal{X}_{1:n})$  almost surely. It then seems reasonable to choose  $a = 1/2$  given that  $\sup_{w \in [0,1]} |\lfloor nw + a \rfloor/n - w| = \max(a, 1-a)/n$  is minimized for  $a = 1/2$  and  $\mathbb{E}(W_{j,u_j}^{\mathbf{x}_{1:n}} \mid \mathcal{X}_{1:n}) = u_j$  almost surely by construction. For this reason, from now on, when using smoothing distributions for which Condition 3.3.3 is not satisfied, instead of  $C_{1:n}^\nu$  in (3.7), we will consider the asymptotically equivalent estimator defined by

$$C_{1:n}^{\nu, \text{cor}}(\mathbf{u}) = \frac{1}{n} \sum_{i=1}^n \mathcal{K}_{\mathbf{R}_i^{1:n} - \mathbf{1}/2}^{\mathbf{x}_{1:n}}(\mathbf{u}), \quad \mathbf{u} \in [0, 1]^d. \quad (3.8)$$

*Remark 3.3.4.* Some additional thinking reveals that (3.8) would have been equivalently obtained if, in (3.5), the classical empirical copula  $C_{1:n}$  in (3.1) were replaced by the empirical d.f. of the modified pseudo-observations  $\tilde{\mathbf{U}}_i^{1:n} = (\mathbf{R}_i^{1:n} - \mathbf{1}/2)/n$ ,  $i \in \{1, \dots, n\}$ . The latter asymptotically equivalent definition of the empirical copula was for instance considered in (Joe, 2015, Section 5.10.1) and has univariate margins that are uniformly closer to the d.f. of the standard uniform distribution than  $C_{1:n}$  in (3.1).

### 3.3.2 General form of the smooth estimators

As already mentioned, we will choose the possibly random smoothing distributions  $\nu_{\mathbf{u}}^{\mathbf{x}_{1:n}}$ ,  $\mathbf{u} \in [0, 1]^d$ , to be either all discrete or all continuous with the understanding that, because of (3.4), if  $u_j \in \{0, 1\}$ , the  $j$ th component of  $\mathbf{W}_{\mathbf{u}}^{\mathbf{x}_{1:n}}$  will be degenerate. In the discrete case, we will further impose Condition 3.3.3 so that, from Proposition 3.3.2, the corresponding smooth empirical copulas of the form (3.7) have standard uniform margins. This property will not hold if the smoothing distributions are chosen continuous. Considering in that case the asymptotically

equivalent definition in (3.8) will however make the uniform distance between the univariate margins of the smooth estimator and the d.f. of the standard uniform distribution smaller than  $1/(2n)$ .

Notice that the expressions of  $C_{1:n}^\nu$  in (3.7) and  $C_{1:n}^{\nu,\text{cor}}$  in (3.8) both depend on the quantity  $\mathcal{K}_r^x$ ,  $\mathbf{x} \in (\mathbb{R}^d)^n$ ,  $\mathbf{r} \in [0, n]^d$ , defined in (3.6). Specifically, recall that, for any  $\mathbf{x} \in (\mathbb{R}^d)^n$ ,  $\mathbf{r} \in [0, n]^d$  and  $\mathbf{u} \in [0, 1]^d$ ,  $\mathcal{K}_r^x(\mathbf{u}) = \mathcal{G}_u^x(\mathbf{r}/n)$ , where  $\mathcal{G}_u^x(\mathbf{w}) = \mathbb{P}(\mathbf{W}_u^x \geq \mathbf{w})$ ,  $\mathbf{w} \in [0, 1]^d$ . For any  $\mathbf{x} \in (\mathbb{R}^d)^n$  and  $\mathbf{u} \in [0, 1]^d$ , let  $\bar{\mathcal{F}}_u^x(\mathbf{w}) = \mathbb{P}(\mathbf{W}_u^x > \mathbf{w})$ ,  $\mathbf{w} \in [0, 1]^d$ , be the survival function of  $\mathbf{W}_u^x$  and note that  $\mathcal{G}_u^x(\mathbf{w}) = \bar{\mathcal{F}}_u^x(\mathbf{w})$ ,  $\mathbf{w} \in [0, 1]^d$ , if  $\mathbf{W}_u^x$  is continuous and that  $\mathcal{G}_u^x(\mathbf{r}/n) = \bar{\mathcal{F}}_u^x\{(\mathbf{r} - \mathbf{1})/n\}$ ,  $\mathbf{r} \in \{1, \dots, n\}^d$ , where  $\mathbf{1} = (1, \dots, 1) \in \mathbb{R}^d$  if  $\mathbf{W}_u^x$  takes its values in the set  $\{0, 1/n, \dots, (n-1)/n, 1\}^d$ . Interestingly enough, for any  $\mathbf{x} \in (\mathbb{R}^d)^n$  and  $\mathbf{u} \in [0, 1]^d$ , we can additionally use the fact that, from Sklar's theorem for survival functions (see, e.g., Hofert et al., 2018, Section 2.5), there exists a copula  $\bar{\mathcal{C}}_u^x$  (called a survival copula of  $\mathbf{W}_u^x$ ) such that

$$\bar{\mathcal{F}}_u^x(\mathbf{w}) = \bar{\mathcal{C}}_u^x\{\bar{\mathcal{F}}_{1,u_1}^x(w_1), \dots, \bar{\mathcal{F}}_{d,u_d}^x(w_d)\}, \quad \mathbf{w} \in [0, 1]^d,$$

where, for any  $j \in \{1, \dots, d\}$ ,  $\bar{\mathcal{F}}_{j,u_j}^x(w) = \mathbb{P}(W_{j,u_j}^x > w)$ ,  $w \in [0, 1]$ , is the  $j$ th univariate margin of  $\bar{\mathcal{F}}_u^x$ . Note that  $\bar{\mathcal{C}}_u^x$  is not uniquely defined unless  $\mathbf{W}_u^x$  is a continuous random vector and that  $\bar{\mathcal{C}}_u^x$  is not a copula of the random vector  $\mathbf{W}_u^x$  but a copula of the random vector  $-\mathbf{W}_u^x$ .

Combining the previous elements, one has that, if the smoothing distributions satisfy Condition 3.3.3, for any  $\mathbf{x} \in (\mathbb{R}^d)^n$ ,  $\mathbf{r} \in [1, n]^d$  and  $\mathbf{u} \in [0, 1]^d$ ,  $\mathcal{K}_r^x(\mathbf{u})$  in (3.6) can be expressed as

$$\mathcal{K}_r^x(\mathbf{u}) = \bar{\mathcal{C}}_u^x[\bar{\mathcal{F}}_{1,u_1}^x\{(r_1 - 1)/n\}, \dots, \bar{\mathcal{F}}_{d,u_d}^x\{(r_d - 1)/n\}], \quad (3.9)$$

which implies that, with probability 1, the smooth empirical copula  $C_{1:n}^\nu$  in (3.7) can be expressed, for any  $\mathbf{u} \in [0, 1]^d$ , as

$$C_{1:n}^\nu(\mathbf{u}) = \frac{1}{n} \sum_{i=1}^n \bar{\mathcal{C}}_u^{\mathbf{x}_{1:n}} \left[ \bar{\mathcal{F}}_{1,u_1}^{\mathbf{x}_{1:n}}\{(R_{i1}^{1:n} - 1)/n\}, \dots, \bar{\mathcal{F}}_{d,u_d}^{\mathbf{x}_{1:n}}\{(R_{id}^{1:n} - 1)/n\} \right], \quad (3.10)$$

whereas, if the smoothing distributions are continuous, as already mentioned, it is better from a marginal perspective to consider the asymptotically equivalent estimator  $C_{1:n}^{\nu,\text{cor}}$  in (3.8) which, with probability 1, can be expressed, for any  $\mathbf{u} \in [0, 1]^d$ , as

$$C_{1:n}^{\nu,\text{cor}}(\mathbf{u}) = \frac{1}{n} \sum_{i=1}^n \bar{\mathcal{C}}_u^{\mathbf{x}_{1:n}} \left[ \bar{\mathcal{F}}_{1,u_1}^{\mathbf{x}_{1:n}}\{(R_{i1}^{1:n} - 1/2)/n\}, \dots, \bar{\mathcal{F}}_{d,u_d}^{\mathbf{x}_{1:n}}\{(R_{id}^{1:n} - 1/2)/n\} \right] \quad (3.11)$$

since, for any  $\mathbf{x} \in (\mathbb{R}^d)^n$ ,  $\mathbf{r} \in [1, n]^d$  and  $\mathbf{u} \in [0, 1]^d$ ,  $\mathcal{K}_{\mathbf{r}-\mathbf{1}/2}^x(\mathbf{u})$  can be expressed as

$$\mathcal{K}_{\mathbf{r}-\mathbf{1}/2}^x(\mathbf{u}) = \bar{\mathcal{C}}_u^x \left[ \bar{\mathcal{F}}_{1,u_1}^x\{(r_1 - 1/2)/n\}, \dots, \bar{\mathcal{F}}_{d,u_d}^x\{(r_d - 1/2)/n\} \right]. \quad (3.12)$$

### 3.3.3 Conditions for being multivariate d.f.s

From a finite-sample perspective, it seems desirable to focus on estimators that are multivariate d.f.s and thus, possibly, genuine copulas. The following conditions are sufficient for that matter.

**Condition 3.3.5** (Condition on the smoothing survival copulas). For any  $\mathbf{x} \in (\mathbb{R}^d)^n$  and  $\mathbf{u} \in [0, 1]^d$ , the copulas  $\bar{C}_{\mathbf{u}}^{\mathbf{x}}$  in (3.9) and (3.12) do not depend on  $\mathbf{u}$ .

**Condition 3.3.6** (Condition on the discrete smoothing survival margins). For any  $\mathbf{x} \in (\mathbb{R}^d)^n$ ,  $j \in \{1, \dots, d\}$  and  $w \in [0, 1)$ , the function  $t \mapsto \bar{\mathcal{F}}_{j,t}^{\mathbf{x}}(w)$  is right-continuous and increasing on  $[0, 1]$ .

The following result is proven in Appendix 3.8.1.

**Proposition 3.3.7** ( $C_{1:n}^{\nu}$  is a multivariate d.f.). Assume that Conditions 3.3.5 and 3.3.6 hold. Then, the smooth empirical copula  $C_{1:n}^{\nu}$  in (3.10) is a multivariate d.f.

**Corollary 3.3.8** ( $C_{1:n}^{\nu}$  is a genuine copula). Assume that Conditions 3.3.1, 3.3.3, 3.3.5 and 3.3.6 hold. Then, the smooth empirical copula  $C_{1:n}^{\nu}$  in (3.5) is a genuine copula.

In the case of continuous smoothing distributions, we consider the following analog of Condition 3.3.6.

**Condition 3.3.9** (Condition on the continuous smoothing survival margins). For any  $\mathbf{x} \in (\mathbb{R}^d)^n$ ,  $j \in \{1, \dots, d\}$  and  $w \in (0, 1)$ , the function  $t \mapsto \bar{\mathcal{F}}_{j,t}^{\mathbf{x}}(w)$  is right-continuous and increasing on  $[0, 1]$ .

The following result is proven in Appendix 3.8.1.

**Proposition 3.3.10** ( $C_{1:n}^{\nu, \text{cor}}$  is a multivariate d.f.). Assume that Conditions 3.3.5 and 3.3.9 hold. Then, the smooth empirical copula  $C_{1:n}^{\nu, \text{cor}}$  in (3.11) is a multivariate d.f.

## 3.4 Estimators based on smoothing distributions with scaled binomial, scaled beta-binomial or beta margins

In order to define specific smooth empirical copulas of the form (3.10) or (3.11) that are multivariate d.f.s, we start by making three proposals for the smoothing survival margins. We first make two proposals for the (discrete) smoothing survival margins of the estimator (3.10) for which Condition 3.3.6 is satisfied and then one proposal for the (continuous) smoothing survival margins of the estimator (3.11) for which Condition 3.3.9 is satisfied. We end this section by discussing the choice of the survival copula of the smoothing distributions.

### 3.4.1 Scaled binomial and beta-binomial smoothing survival margins

Having Corollary 3.3.8 in mind for the estimator (3.10) and given  $u \in [0, 1]$ , we wish to define a random variable  $W$  that takes its values in the set  $\{0, 1/n, \dots, (n-1)/n, 1\}$  such that  $\mathbb{E}(W) = u$ . From (3.4), only the case  $u \in (0, 1)$  actually needs to be dealt with. Following Segers et al. (2017), it is natural to attempt to start from a random variable  $S$  that takes its values in  $\{0, 1, \dots, n\}$  and to set  $W = S/n$ . A first straightforward choice due to Segers et al. (2017) is to take  $S$  to be Binomial( $n, u$ ). As already mentioned, this immediately leads to  $W$  taking its values in the set  $\{0, 1/n, \dots, (n-1)/n, 1\}$  and satisfying  $\mathbb{E}(W) = u$  and  $\text{Var}(W) = u(1-u)/n$ .

As an alternative distribution for the random variable  $S$  taking its values in  $\{0, 1, \dots, n\}$ , we investigate next the possibility of considering the (more dispersed) Beta-Binomial( $n, \alpha, \beta$ ) whose shape parameters  $\alpha > 0$  and  $\beta > 0$  remain to be specified. Recall that the probability mass function of the latter distribution is given by

$$\mathbb{P}(S = s) = \binom{n}{s} \frac{B(s + \alpha, n - s + \beta)}{B(\alpha, \beta)}, \quad s \in \{0, 1, \dots, n\},$$

where  $B$  is the Beta function.

We start from the fact that, since  $S$  is Beta-Binomial( $n, \alpha, \beta$ ),

$$\mathbb{E}(W) = \frac{\alpha}{\alpha + \beta}, \tag{3.13}$$

$$\text{Var}(W) = \frac{\alpha\beta(\alpha + \beta + n)}{n(\alpha + \beta)^2(\alpha + \beta + 1)}. \tag{3.14}$$

Since the expectation of  $W$  is required to be  $u$  by construction, we immediately obtain from (3.13) that

$$\beta = \frac{\alpha(1-u)}{u}, \tag{3.15}$$

which, combined with (3.14), allows us to rewrite the variance of  $W$  as

$$\text{Var}(W) = \frac{u(1-u)(\alpha + un)}{n(\alpha + u)} = \frac{u(1-u)}{n} \times \frac{\alpha + un}{\alpha + u}.$$

The variance of  $W$  is thus the variance of  $W$  if  $S$  were Binomial( $n, u$ ) multiplied by the factor  $(\alpha + un)/(\alpha + u)$ . Ideally, we would want to control how much more dispersed scaled beta-binomial margins are compared to the corresponding scaled binomial margins. To do so, we set the latter factor to be a constant  $\rho > 1$ , that is,

$$\frac{\alpha + un}{\alpha + u} = \rho$$

and attempt to solve for  $\alpha > 0$ . Provided that  $\rho < n$ , we obtain that  $\alpha = u(n - \rho)/(\rho - 1)$  and then, using (3.15), that  $\beta = (1 - u)(n - \rho)/(\rho - 1)$ .

In summary, as an alternative distribution for the discrete random variable  $S$ , we consider the Beta-Binomial( $n, u(n - \rho)/(\rho - 1), (1 - u)(n - \rho)/(\rho - 1)$ ), where  $\rho \in (1, n)$  is an additional parameter. As required, we have that  $\mathbb{E}(W) = u$ . Furthermore,

$$\text{Var}(W) = \rho \times \frac{u(1 - u)}{n}, \quad (3.16)$$

so that the factor  $\rho$  describes how much more dispersed a scaled beta-binomial smoothing survival margin will be compared to the corresponding scaled binomial.

Notice that the probability mass functions of the Binomial( $n, u$ ) and the Binomial( $n, 1 - u$ ) are symmetrical with respect to  $n/2$ , and that, for any  $\rho \in (1, n)$ , the probability mass functions of the Beta-Binomial( $n, u(n - \rho)/(\rho - 1), (1 - u)(n - \rho)/(\rho - 1)$ ) and the Beta-Binomial( $n, (1 - u)(n - \rho)/(\rho - 1), u(n - \rho)/(\rho - 1)$ ) are also symmetrical with respect to  $n/2$ . This implies that, with respect to one coordinate, the smoothing around  $u$  will be the “reflection” of the smoothing around  $1 - u$ .

For any  $u \in [0, 1]$ , let  $\bar{B}_{n,u}$  be the survival function of the Binomial( $n, u$ ) and, for any  $u \in [0, 1]$  and  $\rho \in (1, n)$ , let  $\bar{B}_{n,u,\rho}$  be the survival function of the Beta-Binomial( $n, u(n - \rho)/(\rho - 1), (1 - u)(n - \rho)/(\rho - 1)$ ). Two subclasses of the class of smooth empirical copulas given by (3.10) that, under Condition 3.3.5, still depend on the choice of the possibly data-dependent survival copula  $\bar{C} = \bar{C}^{\mathbf{x}_{1:n}}$ , can thus be defined, for any  $\mathbf{u} \in [0, 1]^d$ , by

$$C_{1:n}^{\bar{B}, \bar{C}}(\mathbf{u}) = \frac{1}{n} \sum_{i=1}^n \bar{C}^{\mathbf{x}_{1:n}} \{ \bar{B}_{n,u_1}(R_{i1}^{1:n} - 1), \dots, \bar{B}_{n,u_d}(R_{id}^{1:n} - 1) \}, \quad (3.17)$$

$$C_{1:n}^{\bar{B}, \bar{C}}(\mathbf{u}) = \frac{1}{n} \sum_{i=1}^n \bar{C}^{\mathbf{x}_{1:n}} \{ \bar{B}_{n,u_1,\rho}(R_{i1}^{1:n} - 1), \dots, \bar{B}_{n,u_d,\rho}(R_{id}^{1:n} - 1) \}, \quad (3.18)$$

respectively. It is of course possible to imagine a version of the second estimator for which the common additional parameter  $\rho$  of the  $d$  scaled beta-binomial smoothing survival margins depends on the data, that is,  $\rho = \rho^{\mathbf{x}_{1:n}}$  (subject to the constraint that  $\rho^{\mathbf{x}_{1:n}} \in (1, n)$  almost surely). By construction, the smoothing survival margins satisfy Condition 3.3.3, which, according to Proposition 3.3.2, implies that  $C_{1:n}^{\bar{B}, \bar{C}}$  and  $C_{1:n}^{\bar{B}, \bar{C}}$  have standard uniform univariate margins. As far as Condition 3.3.6 is concerned, we have the following results proven in Appendix 3.8.2.

**Proposition 3.4.1.** *Condition 3.3.6 is satisfied with  $\bar{\mathcal{F}}_{j,t}^{\mathbf{x}}$  defined by  $\bar{\mathcal{F}}_{j,t}^{\mathbf{x}}(w) = \bar{B}_{n,t}(nw)$ ,  $w \in [0, 1]$ . Specifically, for any  $w \in [0, 1]$ , the function  $t \mapsto \bar{B}_{n,t}(nw)$  is continuous and increasing on  $[0, 1]$ .*

**Proposition 3.4.2.** *Condition 3.3.6 is satisfied with  $\bar{\mathcal{F}}_{j,t}^{\mathbf{x}}$  defined by  $\bar{\mathcal{F}}_{j,t}^{\mathbf{x}}(w) = \bar{B}_{n,t,\rho}(nw)$ ,  $w \in [0, 1]$ , for any  $\rho \in (1, n)$ . Specifically, for any  $\rho \in (1, n)$  and  $w \in [0, 1]$ , the function  $t \mapsto \bar{B}_{n,t,\rho}(nw)$  is continuous and increasing on  $[0, 1]$ .*

Hence, according to Proposition 3.3.7, the estimators  $C_{1:n}^{\bar{B}, \bar{C}}$  and  $C_{1:n}^{\bar{B}, \bar{C}}$  are multivariate d.f.s and therefore, according to Corollary 3.3.8, genuine copulas if Condition 3.3.1 holds. Notice also

that the smooth empirical copula  $C_{1:n}^{\bar{B},\bar{c}}$  coincides with the empirical beta copula  $C_{1:n}^{\text{beta}}$  in (3.2) if  $\bar{C}^{\mathbf{x}_{1:n}}$  is taken to be the independence copula  $\Pi$ . Based on (3.17), an alternative notation for the empirical beta copula is thus  $C_{1:n}^{\bar{B},\Pi}$ .

*Remark 3.4.3.* Let  $m \in \mathbb{N}$  and recall from Remark 3.2.4 that if, for any  $\mathbf{x} \in (\mathbb{R}^d)^n$  and  $\mathbf{u} \in [0, 1]^d$ ,  $\nu_{\mathbf{u}}^{\mathbf{x}}$  is the law of the random vector  $(S_{m,1,u_1}/m, \dots, S_{m,d,u_d}/m)$  where  $S_{m,1,u_1}, \dots, S_{m,d,u_d}$  are independent random variables such that, for each  $j \in \{1, \dots, d\}$ ,  $S_{m,j,u_j}$  is Binomial( $m, u_j$ ), then, from Section 3 in Segers et al. (2017),  $C_{1:n}^{\nu}$  in (3.5) is the empirical Bernstein copula with polynomial degrees all equal to  $m$ . It immediately follows that the expectation and variance of the  $j$ th margin of the underlying smoothing distribution corresponding to  $\mathbf{u} \in [0, 1]^d$  are  $u_j$  and  $u_j(1 - u_j)/m$ , respectively. Let  $\rho \in (1, n)$  and let us focus on the special case  $m = \lceil n/\rho \rceil$ . The underlying smoothing distribution corresponding to  $\mathbf{u} \in [0, 1]^d$  can then be regarded as close to the smoothing distribution corresponding to the same  $\mathbf{u}$  involved in the definition of the estimator  $C_{1:n}^{\bar{B},\Pi}$  generically defined in (3.18) with dispersion parameter  $\rho$ . Indeed, both smoothing distributions have independent copula, expectation  $\mathbf{u}$  and the variances of their  $j$ th margin are  $u_j(1 - u_j)/\lceil n/\rho \rceil$  and  $\rho u_j(1 - u_j)/n$ , respectively (see (3.16)). The estimator  $C_{1:n}^{\bar{B},\Pi}$  with dispersion parameter  $\rho$  can however be regarded as an advantageous replacement of the empirical Bernstein copula with polynomial degrees all equal to  $\lceil n/\rho \rceil$ . Indeed, if Condition 3.3.1 holds, the former is a genuine copula by Corollary 3.3.8 (no matter how  $\rho \in (1, n)$  is chosen) whereas the latter is a genuine copula only when  $n$  is a multiple of  $\lceil n/\rho \rceil$  by Proposition 2.5 in Segers et al. (2017). When  $\rho = 3$ , the disadvantages in terms of finite-sample performance of the empirical Bernstein copula with polynomial degrees all equal to  $\lceil n/\rho \rceil$  are highlighted in Fig. 3 of Segers et al. (2017).

### 3.4.2 Beta smoothing survival margins

Since the support of the beta distribution is included in  $[0, 1]$ , it is natural to attempt to use it to define the continuous smoothing survival margins involved in the definition of the estimator (3.11). Thus, given  $u \in [0, 1]$ , let  $W$  be a random variable with the beta distribution Beta( $\alpha, \beta$ ) whose shape parameters  $\alpha > 0$  and  $\beta > 0$  need to be determined so that  $\mathbb{E}(W) = u$ . Again, from (3.4), only the case  $u \in (0, 1)$  needs to be addressed. We know that

$$\mathbb{E}(W) = \frac{\alpha}{\alpha + \beta}, \tag{3.19}$$

$$\text{Var}(W) = \frac{\alpha\beta}{(\alpha + \beta)^2(\alpha + \beta + 1)}. \tag{3.20}$$

From the fact that the expectation of  $W$  is required to be  $u$  and (3.19), we obtain that  $\beta = \alpha(1 - u)/u$ , which, combined with (3.20), implies in turn that

$$\text{Var}(W) = \frac{u^2(1 - u)}{\alpha + u} = \frac{u(1 - u)}{n} \times \frac{nu}{\alpha + u}.$$

Proceeding as for the scaled beta-binomial smoothing survival margins but keeping in mind that the variance of the beta distribution with expectation  $u$  can be arbitrarily small, we set

$$\frac{nu}{\alpha + u} = \rho > 0$$

and attempt to solve for  $\alpha > 0$ . Provided that  $\rho < n$ , we obtain that  $\alpha = u(n - \rho)/\rho$  and, consequently, that  $\beta = (1 - u)(n - \rho)/\rho$ . The distribution of  $W$  can thus be taken to be the  $\text{Beta}(u(n - \rho)/\rho, (1 - u)(n - \rho)/\rho)$ , where  $\rho \in (0, n)$  is an additional parameter. As required, we have that  $\mathbb{E}(W) = u$ , whereas the parameter  $\rho$  controls the variance of  $W$  which is given by

$$\text{Var}(W) = \rho \times \frac{u(1 - u)}{n}. \quad (3.21)$$

Hence, the variance of  $W$  can again be compared to the variance of the corresponding scaled binomial survival margin. Notice however that, unlike for scaled beta-binomial survival margins, beta survival margins can be underdispersed compared to the corresponding scaled binomial survival margins.

From the fact that the densities of the  $\text{Beta}(\alpha, \beta)$  and the  $\text{Beta}(\beta, \alpha)$  are symmetrical with respect to  $1/2$ , we again have that, with respect to one coordinate, the smoothing around  $u$  will be the ‘‘reflection’’ of the smoothing around  $1 - u$ .

For any  $u \in [0, 1]$  and  $\rho \in (0, n)$ , let  $\bar{\beta}_{n,u,\rho}$  be the survival function of the  $\text{Beta}(u(n - \rho)/\rho, (1 - u)(n - \rho)/\rho)$ ,  $\rho \in (0, n)$ . One subclass of the class of smooth empirical copulas (3.11) that, under Condition 3.3.5, still depends on the choice of the possibly data-dependent survival copula  $\bar{C} = \bar{C}^{\mathbf{x}_{1:n}}$ , can then be defined by

$$C_{1:n}^{\bar{\beta}, \bar{C}}(\mathbf{u}) = \frac{1}{n} \sum_{i=1}^n \bar{C}^{\mathbf{x}_{1:n}} [\bar{\beta}_{n,u_1,\rho}\{(R_{i1}^{1:n} - 1/2)/n\}, \dots, \bar{\beta}_{n,u_d,\rho}\{(R_{id}^{1:n} - 1/2)/n\}], \quad \mathbf{u} \in [0, 1]^d. \quad (3.22)$$

As for the estimator (3.18), it is possible to imagine a version of (3.22) for which the common additional parameter  $\rho$  of the  $d$  beta smoothing survival margins depends on the data, that is,  $\rho = \rho^{\mathbf{x}_{1:n}}$  (subject to the constraint that  $\rho^{\mathbf{x}_{1:n}} \in (0, n)$  almost surely). As far as Condition 3.3.9 is concerned, we have the following result proven in Appendix 3.8.2.

**Proposition 3.4.4.** *Condition 3.3.9 is satisfied with  $\bar{\mathcal{F}}_{j,t}^{\mathbf{x}} = \bar{\beta}_{n,t,\rho}$ , for any  $\rho \in (0, n)$ . Specifically, for any  $\rho \in (0, n)$  and  $w \in (0, 1)$ , the function  $t \mapsto \bar{\beta}_{n,t,\rho}(w)$  is continuous and increasing on  $[0, 1]$ .*

Hence, according to Proposition 3.3.10,  $C_{1:n}^{\bar{\beta}, \bar{C}}$  is a multivariate d.f. If Condition 3.3.1 holds, it is however not a genuine copula since it does not have standard uniform margins. The latter can for instance be verified numerically using the R implementation of the estimator provided on the web page of the first author.



### 3.4.3 On the survival copula of the smoothing distributions

In view of Propositions 3.3.7 and 3.3.10, it seems desirable to choose the survival copulas of the smoothing distributions appearing in (3.10) and (3.11) such that Condition 3.3.5 holds. In order to propose a meaningful choice for the resulting possibly data-dependent survival copula  $\bar{C}^{\mathbf{x}_{1:n}}$ , we first gather hereafter some theoretical facts.

Assume that Condition 3.3.1 holds and fix  $\mathbf{u} \in [0, 1]^d$ . If  $\mathbf{W}_{\mathbf{u}}^{\mathbf{x}_{1:n}}$  takes its values in the set  $\{0, 1/n, \dots, (n-1)/n, 1\}^d$ , it is easy to verify from (3.10) that

$$\mathbb{E}\{C_{1:n}^{\nu}(\mathbf{u})\} = \mathbb{E}\left(\bar{C}^{\mathbf{x}_{1:n}}[\bar{\mathcal{F}}_{1,u_1}^{\mathbf{x}_{1:n}}\{(R_{11}^{1:n} - 1)/n\}, \dots, \bar{\mathcal{F}}_{d,u_d}^{\mathbf{x}_{1:n}}\{(R_{1d}^{1:n} - 1)/n\}]\right) \quad (3.23)$$

since  $\mathbf{X}_1, \dots, \mathbf{X}_n$  are identically distributed. Note furthermore that, for any  $j \in \{1, \dots, d\}$ ,

$$\begin{aligned} \mathbb{E}[\bar{\mathcal{F}}_{j,u_j}^{\mathbf{x}_{1:n}}\{(R_{1j}^{1:n} - 1)/n\}] &= \mathbb{E}\left(\mathbb{E}\left[\bar{\mathcal{F}}_{j,u_j}^{\mathbf{x}_{1:n}}\{(R_{1j}^{1:n} - 1)/n\} \mid \mathbf{x}_{1:n}\right]\right) = \mathbb{E}\left[\frac{1}{n} \sum_{i=1}^n \bar{\mathcal{F}}_{j,u_j}^{\mathbf{x}_{1:n}}\{(i-1)/n\}\right] \\ &= \mathbb{E}\left[\frac{1}{n} \sum_{i=1}^n \mathbb{P}\{W_{j,u_j}^{\mathbf{x}_{1:n}} > (i-1)/n \mid \mathbf{x}_{1:n}\}\right] = \mathbb{E}\left[\frac{1}{n} \sum_{i=1}^n \mathbb{E}\left\{\mathbf{1}\{W_{j,u_j}^{\mathbf{x}_{1:n}} \geq i/n\} \mid \mathbf{x}_{1:n}\right\}\right] \\ &= \mathbb{E}\left[\mathbb{E}\left\{\frac{1}{n} \sum_{i=1}^n \mathbf{1}\{i \leq nW_{j,u_j}^{\mathbf{x}_{1:n}}\} \mid \mathbf{x}_{1:n}\right\}\right] = \mathbb{E}\left\{\mathbb{E}\left(\frac{\lfloor nW_{j,u_j}^{\mathbf{x}_{1:n}} \rfloor}{n} \mid \mathbf{x}_{1:n}\right)\right\} \\ &= \mathbb{E}\left\{\mathbb{E}(W_{j,u_j}^{\mathbf{x}_{1:n}} \mid \mathbf{x}_{1:n})\right\} = \mathbb{E}(u_j) = u_j. \end{aligned} \quad (3.24)$$

In other words, the expected value of  $C_{1:n}^{\nu}(\mathbf{u})$  is the expectation of a random variable obtained by applying the copula  $\bar{C}^{\mathbf{x}_{1:n}}$  to the components of a random vector with expectation  $\mathbf{u}$ . Similarly, if  $\mathbf{W}_{\mathbf{u}}^{\mathbf{x}_{1:n}}$  is continuous, from (3.11), we have that

$$\mathbb{E}\{C_{1:n}^{\nu, \text{COR}}(\mathbf{u})\} = \mathbb{E}\left(\bar{C}^{\mathbf{x}_{1:n}}[\bar{\mathcal{F}}_{1,u_1}^{\mathbf{x}_{1:n}}\{(R_{11}^{1:n} - 1/2)/n\}, \dots, \bar{\mathcal{F}}_{d,u_d}^{\mathbf{x}_{1:n}}\{(R_{1d}^{1:n} - 1/2)/n\}]\right) \quad (3.25)$$

and, moreover, proceeding similarly to (3.24), that

$$\mathbb{E}[\bar{\mathcal{F}}_{j,u_j}^{\mathbf{x}_{1:n}}\{(R_{1j}^{1:n} - 1/2)/n\}] = \mathbb{E}\left(\frac{\lfloor nW_{j,u_j}^{\mathbf{x}_{1:n}} + 1/2 \rfloor}{n}\right), \quad (3.26)$$

which, combined with the fact that  $\sup_{w \in [0,1]} |\lfloor nw + 1/2 \rfloor/n - w| = 1/(2n)$ , implies that

$$|\mathbb{E}[\bar{\mathcal{F}}_{j,u_j}^{\mathbf{x}_{1:n}}\{(R_{1j}^{1:n} - 1/2)/n\}] - u_j| \leq 1/(2n). \quad (3.27)$$

In other words, in the case of a continuous smoothing distribution, the expected value of  $C_{1:n}^{\nu, \text{COR}}(\mathbf{u})$  is the expectation of a random variable obtained by applying the copula  $\bar{C}^{\mathbf{x}_{1:n}}$  to the components of a random vector whose expectation is within  $1/(2n)$  of  $\mathbf{u}$ .

Given that  $C_{1:n}^{\nu}(\mathbf{u})$  and  $C_{1:n}^{\nu, \text{COR}}(\mathbf{u})$  are estimators of  $C(\mathbf{u})$ , the previous derivations suggest that it may be meaningful to expect that, in many situations, their biases will be minimized (or at

least will be small) if the survival copula  $\bar{C}^{\mathbf{x}_{1:n}}$  of  $\mathbf{W}_u^{\mathbf{x}_{1:n}}$  is taken equal to the copula  $C$ . When  $C$  is the independence copula  $\Pi$ , the above choice is optimal if Condition 3.3.3 holds. Indeed, setting  $\bar{C}^{\mathbf{x}_{1:n}} = \Pi$  in (3.23) and using (3.24) immediately yields that  $C_{1:n}'(\mathbf{u})$  is an unbiased estimator of  $C(\mathbf{u}) = \Pi(\mathbf{u})$ . Similarly, if  $\mathbf{W}_u^{\mathbf{x}_{1:n}}$  is continuous, one straightforwardly obtains from (3.25), (3.26) and (3.27) that the bias of  $C_{1:n}'^{\text{cor}}(\mathbf{u})$  will decrease quickly as  $n$  increases.

The previous discussion focusing on the bias thus suggests, if  $C$  were known, to choose the survival copula  $\bar{C}^{\mathbf{x}_{1:n}}$  in (3.10) and (3.11) equal to  $C$ . Since  $C$  is not known, a natural solution consists of taking  $\bar{C}^{\mathbf{x}_{1:n}}$  equal to a pilot estimate of  $C$  based on  $\mathbf{x}_{1:n}$ . This aspect will be empirically investigated in Section 3.5.

### 3.5 Suggested data-adaptive alternatives to the empirical beta copula

The first aim of this section is to study, under Condition 3.3.5, the finite-sample performance of the estimators  $C_{1:n}^{\bar{B},\bar{C}}$  in (3.17),  $C_{1:n}^{\bar{B},\bar{C}}$  in (3.18) and  $C_{1:n}^{\bar{B},\bar{C}}$  in (3.22) in the case when the survival copula  $\bar{C}^{\mathbf{x}_{1:n}}$  does not depend on the data. These preliminary investigations will then be used to propose a natural way of choosing  $\bar{C}^{\mathbf{x}_{1:n}}$  from the data in the expressions of  $C_{1:n}^{\bar{B},\bar{C}}$ ,  $C_{1:n}^{\bar{B},\bar{C}}$  and  $C_{1:n}^{\bar{B},\bar{C}}$ . The finite-sample performance of the resulting estimators will be finally studied for many bivariate and trivariate data-generating models.

#### 3.5.1 Experiments for understanding the influence of the spread and the shape of the smoothing distributions

Following Segers et al. (2017), for each sample size  $n$  under consideration, each data generating copula  $C$  and each copula estimator  $\hat{C}_{1:n}$  under investigation, we estimated the following three performance measures:

$$\begin{aligned} \text{integrated squared bias: } & \int_{[0,1]^d} \left[ \mathbb{E} \left\{ \hat{C}_{1:n}(\mathbf{u}) - C(\mathbf{u}) \right\} \right]^2 d\mathbf{u}, \\ \text{integrated variance: } & \int_{[0,1]^d} \mathbb{E} \left( \left[ \hat{C}_{1:n}(\mathbf{u}) - \mathbb{E}\{\hat{C}_{1:n}(\mathbf{u})\} \right]^2 \right) d\mathbf{u}, \\ \text{integrated mean squared error: } & \int_{[0,1]^d} \mathbb{E} \left[ \left\{ \hat{C}_{1:n}(\mathbf{u}) - C(\mathbf{u}) \right\}^2 \right] d\mathbf{u}. \end{aligned}$$

To do so, we proceeded by Monte Carlo simulation using 20,000 independent random samples of size  $n$  from  $C$  and applied the trick described in detail in Appendix B of Segers et al. (2017). All the numerical experiments were carried out using the R statistical environment and its packages `copula` Hofert et al. (2022) and `extraDistr` Wolodzko (2020).

As a first experiment, we considered  $n \in \{10, 30, 50\}$  and  $C$  to be the bivariate Frank copula Genest (1987) with a Kendall's tau in  $\{0, -0.3, -0.6, -0.9\}$ . We then estimated the above three

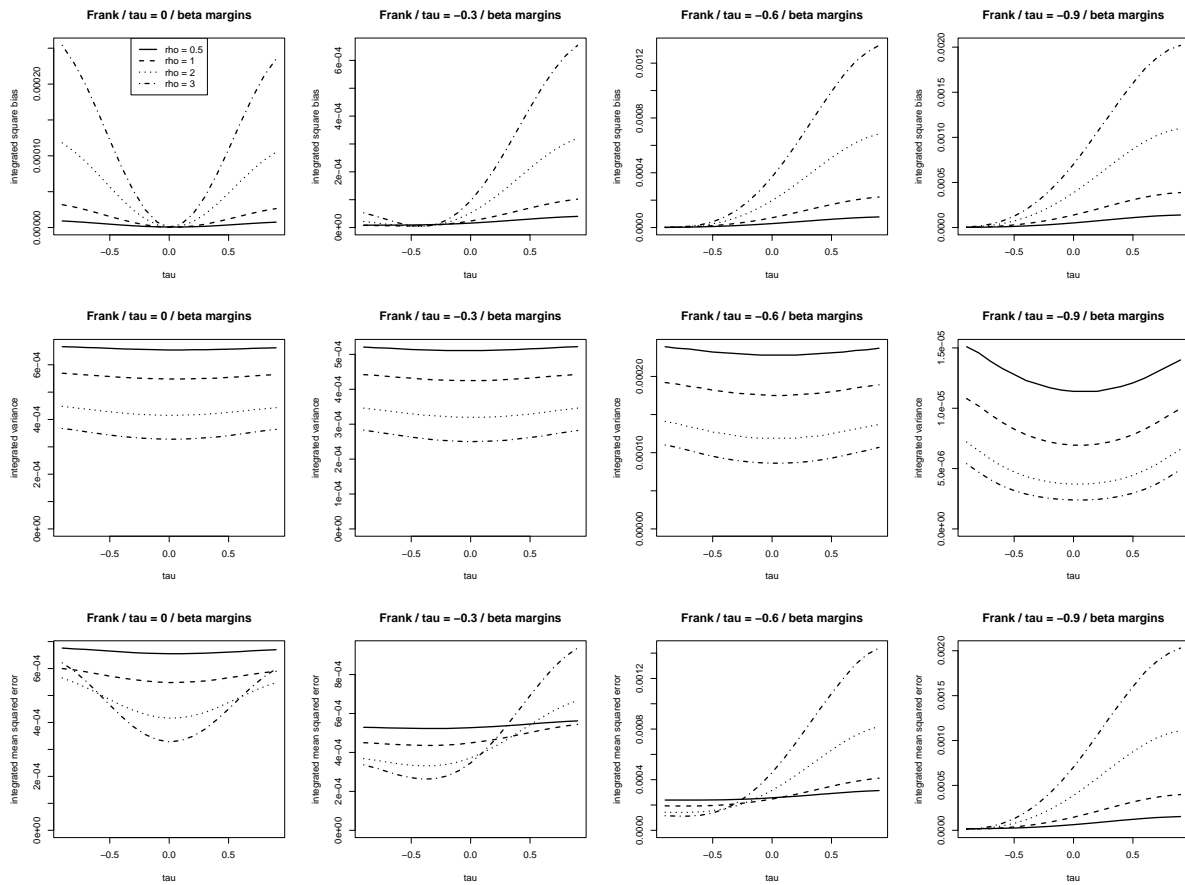


FIGURE 3.1: Influence of the dispersion parameter  $\rho$  of the beta smoothing survival margins and of the value of Kendall's tau (given on the  $x$ -axis of each panel) of the smoothing Frank copula  $\bar{C}^{\mathbf{x}_{1:n}}$  (not taken to depend on the data) on the three performances measures of the smooth estimator  $C_{1:n}^{\bar{\beta}, \bar{C}}$  in (3.22) computed from samples of size  $n = 30$  from a Frank copula with a Kendall's tau in  $\{0, -0.3, -0.6, -0.9\}$ .

performance measures for the smooth empirical copula  $C_{1:n}^{\bar{\beta}, \bar{C}}$  in (3.22) with  $\rho \in \{0.5, 1, 2, 3\}$  and  $\bar{C}^{\mathbf{x}_{1:n}}$  the Frank copula with a Kendall's tau in  $\{-0.9, -0.8, \dots, 0.9\}$  (thus not taken to depend on the data). The results are represented in Fig. 3.1 for  $n = 30$  and are qualitatively similar for  $n \in \{10, 50\}$ . Each column of graphs corresponds to a different Kendall's tau for the data generating Frank copula. The first (resp. second, third) row of plots reports the integrated squared bias (resp. integrated variance, integrated mean squared error). The four curves in each panel represent the four considered values of the dispersion parameter  $\rho$  and give the value of the performance measure against the value of Kendall's tau of the (data-independent) Frank copula  $\bar{C}^{\mathbf{x}_{1:n}}$  of the smoothing survival distribution.

From the first row of graphs in Fig. 3.1, we see that, as expected from the discussion in Section 3.4.3, the integrated squared bias is minimized when the survival copula  $\bar{C}^{\mathbf{x}_{1:n}}$  is close to the data generating copula  $C$ . The larger the value of the marginal dispersion parameter  $\rho$ , the more obvious the previous conclusion is. The second row of plots reveals, on one hand, that the shape of the smoothing distribution controlled by  $\bar{C}^{\mathbf{x}_{1:n}}$  has relatively little influence on the

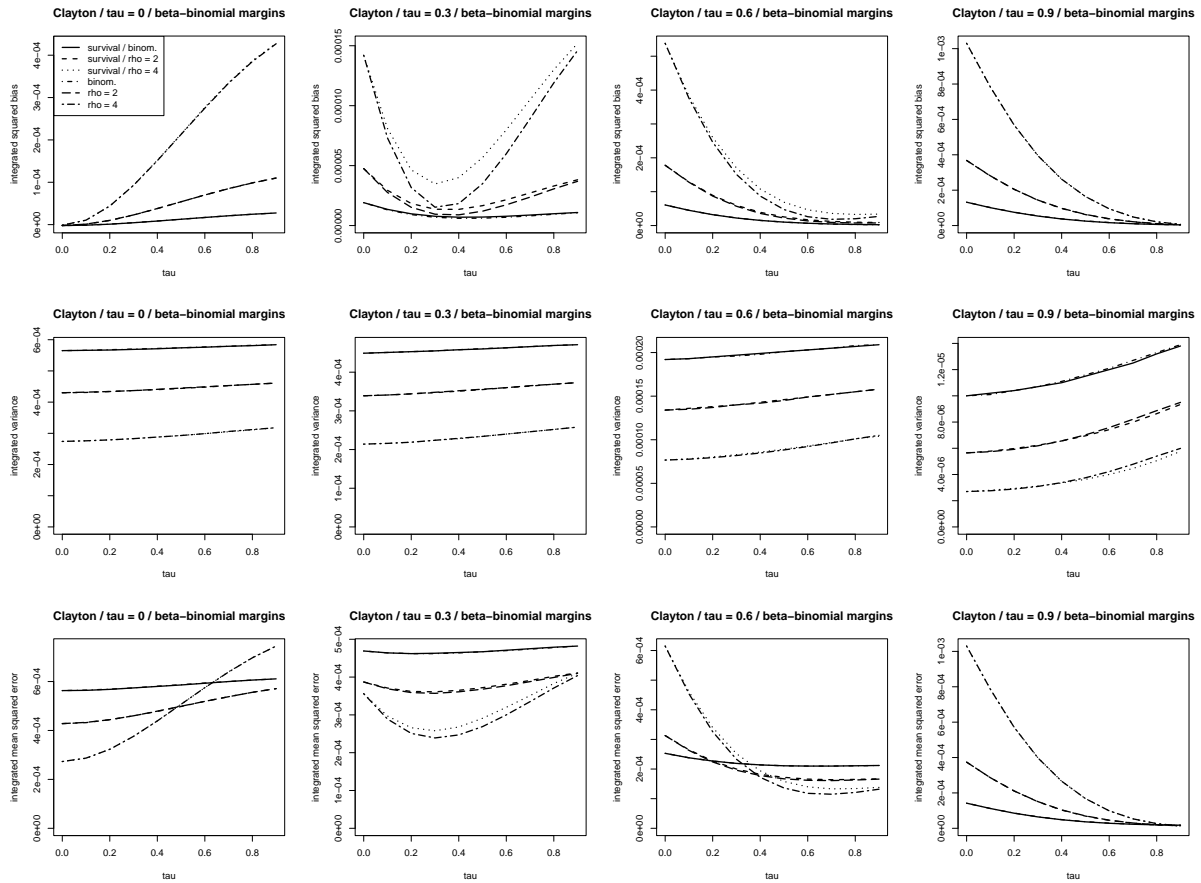


FIGURE 3.2: Influence of the dispersion parameter  $\rho$  of the scaled beta-binomial smoothing survival margins and of the value of Kendall's tau (given on the  $x$ -axis of each panel) of the smoothing Clayton or survival Clayton copula  $\bar{C}^{\mathcal{X}_{1:n}}$  (not taken to depend on the data) on the three performances measures of the smooth estimators  $C_{1:n}^{\bar{B}, \bar{C}}$  in (3.17) and  $C_{1:n}^{\bar{B}, \bar{C}}$  in (3.18) computed from samples of size  $n = 30$  drawn from a Clayton copula with a Kendall's tau in  $\{0, 0.3, 0.6, 0.9\}$ .

integrated variance (although the latter seems always minimized when  $\bar{C}^{\mathcal{X}_{1:n}}$  is close to the independence copula  $\Pi$ ) and, on the other hand, that the larger  $\rho$ , the lower the integrated variance. The third row of graphs combines the previous conclusions and shows that the integrated mean squared error seems minimized when  $\bar{C}^{\mathcal{X}_{1:n}}$  is close to  $C$  and  $\rho$  is large provided the absolute value of Kendall's tau of  $C$  is strictly smaller than 0.9. In the case of strongly (negatively) dependent observations, the lowest integrated mean squared error seems also reached for  $\rho = 0.5$ , the latter setting having the advantage of being hardly unaffected by the choice of  $\bar{C}^{\mathcal{X}_{1:n}}$ .

As a second experiment, we considered a similar setting but with  $C$  a bivariate Clayton copula Clayton (1978) and for the estimators  $C_{1:n}^{\bar{B}, \bar{C}}$  in (3.17) and  $C_{1:n}^{\bar{B}, \bar{C}}$  in (3.18) involving scaled binomial and scaled beta-binomial smoothing survival margins with dispersion parameter  $\rho \in \{2, 4\}$ , respectively. This time, we also allowed  $\bar{C}^{\mathcal{X}_{1:n}}$  to be a survival Clayton copula (still not taken to depend on the data). Note that a Clayton copula and the corresponding survival copula have the same value of Kendall's tau whereas they are very different: the former is lower tail

dependent while the latter is upper tail dependent (see, e.g., Hofert et al., 2018, Section 3.4.1). The results are represented in Fig. 3.2 for  $n = 30$ . The conclusions are the same as those drawn after the first experiment. In particular, the curves for  $\bar{C}^{\mathbf{x}_{1:n}}$  the Clayton copula and  $\bar{C}^{\mathbf{x}_{1:n}}$  the corresponding survival Clayton copula for the same value of  $\rho$  in Fig. 3.2 confirm that the integrated bias and the integrated mean squared error are minimized when  $\bar{C}^{\mathbf{x}_{1:n}}$  is approximately equal to  $C$ . The conclusions remain qualitatively the same for  $n \in \{10, 50\}$  or when the Clayton family is replaced by the Gumbel–Hougaard family (Gumbel, 1961; Hougaard, 1986).

### 3.5.2 Suggested data-adaptive smooth estimators and their finite-sample performance

The previous experiments confirm what the theoretical hints given in Section 3.4.3 already suggested: it seems meaningful to take the survival copula  $\bar{C}^{\mathbf{x}_{1:n}}$  of the smoothing distributions equal to the data generating copula  $C$ . Since the latter is unknown and since the empirical beta copula  $C_{1:n}^{\text{beta}}$  in (3.2) is probably one of the best available estimators of  $C$ , it is a natural choice for  $\bar{C}^{\mathbf{x}_{1:n}}$ . As far as the dispersion parameter  $\rho$  is concerned, the results reported in Section 3.5.1 suggest that taking its value to be larger than one, say  $\rho \in \{2, 4\}$ , may be a good general choice. Two suggested smooth data-adaptive copula estimators are thus  $C_{1:n}^{\bar{\beta}, C_{1:n}^{\text{beta}}}$  and  $C_{1:n}^{\bar{\beta}, C_{1:n}^{\text{beta}}}$  with  $\rho \in \{2, 4\}$ . The first one, generically defined in (3.18), has scaled beta-binomial smoothing survival margins and uses the empirical beta copula for  $\bar{C}^{\mathbf{x}_{1:n}}$ . The second one, generically defined in (3.22), has beta smoothing survival margins and also uses the empirical beta copula for  $\bar{C}^{\mathbf{x}_{1:n}}$ . As competitors to these estimators in our Monte Carlo experiments, we considered:

- the empirical beta copula  $C_{1:n}^{\text{beta}}$  (which is the same as the estimator  $C_{1:n}^{\bar{\beta}, \Pi}$  in (3.17) with  $\bar{C}^{\mathbf{x}_{1:n}} = \Pi$ ),
- the estimators  $C_{1:n}^{\bar{\beta}, \Pi}$  and  $C_{1:n}^{\rho, \bar{\beta}, \Pi}$  with  $\rho \in \{2, 4\}$ , which can be regarded as smoother versions of  $C_{1:n}^{\text{beta}} = C_{1:n}^{\bar{\beta}, \Pi}$ ,
- and the estimators  $C_{1:n}^{\bar{\beta}, \Pi}$  and  $C_{1:n}^{\rho, \bar{\beta}, C_{1:n}^{\text{beta}}}$  with  $\rho = 0.5$  which are marginally rougher than all of the previously considered ones.

In a first series of experiments, we considered  $d \in \{2, 3\}$  and  $C$  to be either the  $d$ -dimensional Clayton, Gumbel–Hougaard or Frank copula whose bivariate margins have a Kendall’s tau of  $\tau$  and we estimated the integrated mean squared error of the aforementioned estimators for  $n \in \{10, 20, 30, 40\}$  and for  $\tau$  in either  $\{0, 0.1, \dots, 0.9\}$  or in  $\{-0.9, -0.8, \dots, 0\}$ . The estimated integrated mean squared errors are given against Kendall’s tau in the panels of Fig. 3.3 in which the following abbreviations are used:

- “e.b.c.” for “empirical beta copula”,

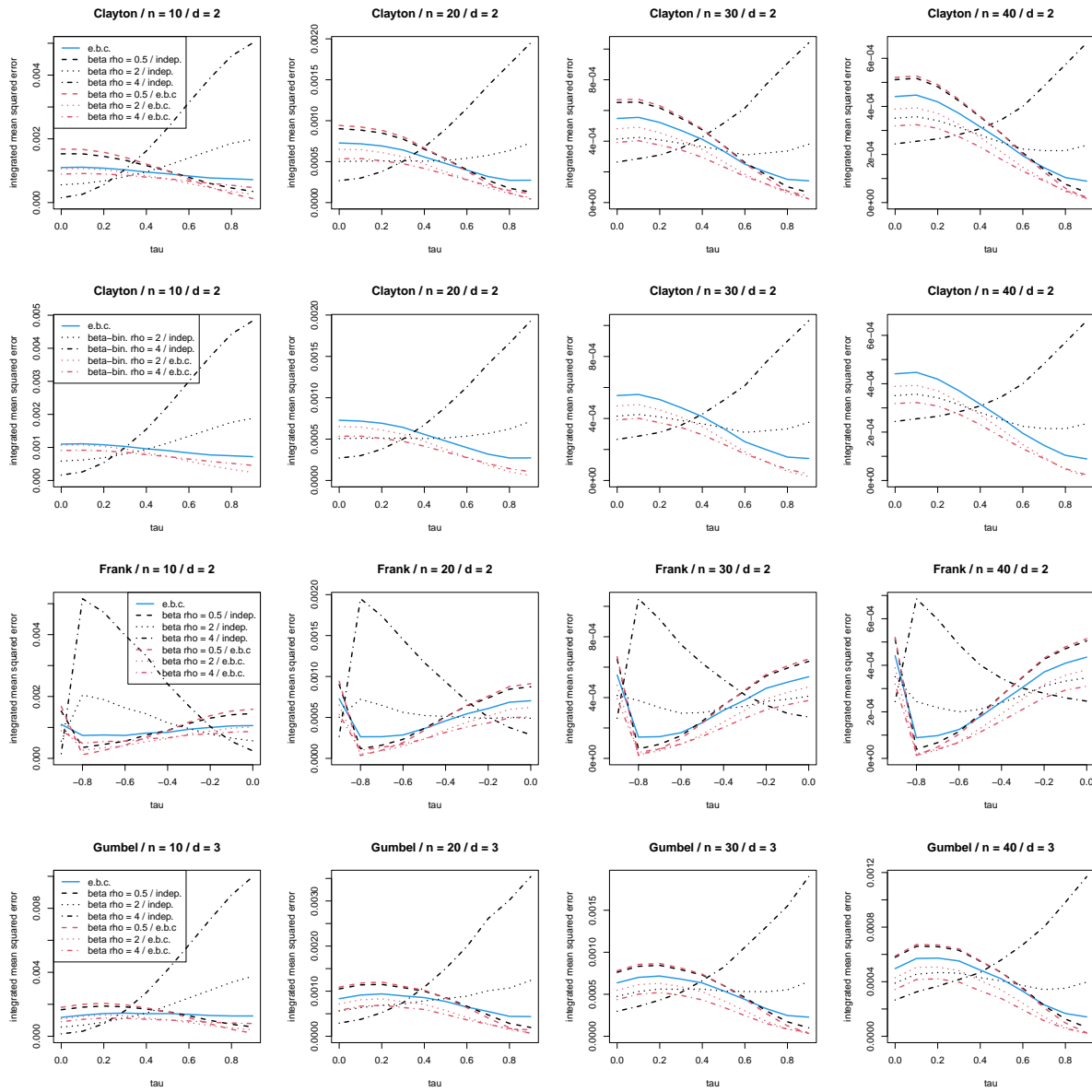


FIGURE 3.3: Estimated integrated mean squared errors against Kendall's tau  $\tau$  for  $\tau$  in either  $\{0, 0.1, \dots, 0.9\}$  or  $\{-0.9, -0.8, \dots, 0\}$  of various estimators of  $C$  for  $C$  the bivariate Clayton (first and second row of graphs), the bivariate Frank (third row of graphs) or the trivariate Gumbel–Hougaard copula (fourth row of graphs) for  $n \in \{10, 20, 30, 40\}$ . The abbreviations used in the legends are those defined in Section 3.5.2.

- “beta-bin. rho = 2 / indep” for the estimator  $C_{1:n}^{\bar{\beta}, \Pi}$  with  $\rho = 2$ ,
- “beta-bin. rho = 2 / e.b.c.” for the estimator  $C_{1:n}^{\bar{\beta}, C_{1:n}^{\text{beta}}}$  with  $\rho = 2$ ,
- “beta rho = 2 / indep” for the estimator  $C_{1:n}^{\bar{\beta}, \Pi}$  with  $\rho = 2$ ,
- “beta rho = 2 / e.b.c.” for the estimator  $C_{1:n}^{\bar{\beta}, C_{1:n}^{\text{beta}}}$  with  $\rho = 2$ ,

and similarly for the estimators with  $\rho = 4$ .

The following conclusions can be drawn from Fig. 3.3:

- the estimators  $C_{1:n}^{\bar{\beta}, C_{1:n}^{\text{beta}}}$  and  $C_{1:n}^{\bar{\beta}, C_{1:n}^{\text{beta}}}$  with  $\rho \in \{2, 4\}$  are uniformly better than the empirical beta copula in terms of integrated mean squared error for the data generating models under consideration; setting  $\rho$  to 4 is overall better than setting  $\rho = 2$ , except when, approximately,  $n \leq 30$  and  $|\tau| \geq 0.6$ ,
- as expected from Section 3.4.3 and the experiments reported in Section 3.5.1, all other settings being equal, the estimators for which  $\bar{C}^{\mathbf{x}}_{1:n} = \Pi$  are better than their analogs for which  $\bar{C}^{\mathbf{x}}_{1:n} = C_{1:n}^{\text{beta}}$  when  $\tau$  is close to 0; their advantage, particularly large for small  $n$ , decreases as  $n$  increases, that is, as the performance of the pilot estimator  $C_{1:n}^{\text{beta}}$  improves,
- as expected from the experiments reported in Section 3.5.1, setting  $\rho = 0.5$  in the expressions of  $C_{1:n}^{\bar{\beta}, C_{1:n}^{\text{beta}}}$  and  $C_{1:n}^{\bar{\beta}, \Pi}$  can have advantages only when  $|\tau|$  is large and  $n$  is small,
- as can be seen from the first and second rows of plots (and from other non reported results), for the same value of  $\rho > 1$ , there is hardly any difference between the estimators  $C_{1:n}^{\bar{\beta}, C_{1:n}^{\text{beta}}}$  and  $C_{1:n}^{\bar{\beta}, C_{1:n}^{\text{beta}}}$  (resp.  $C_{1:n}^{\bar{\beta}, \Pi}$  and  $C_{1:n}^{\bar{\beta}, \Pi}$ ) in terms of integrated mean squared error.

Note that we have no explanations for the somehow surprising behavior occurring in the third line of graphs when  $\tau$  changes from  $-0.8$  to  $-0.9$ .

In a last series of experiments, we compared the integrated mean squared error of  $C_{1:n}^{\text{beta}}$  with that of  $C_{1:n}^{\bar{\beta}, C_{1:n}^{\text{beta}}}$ ,  $C_{1:n}^{\bar{\beta}, \Pi}$ ,  $C_{1:n}^{\bar{\beta}, C_{1:n}^{\text{beta}}}$  and  $C_{1:n}^{\bar{\beta}, \Pi}$  for  $\rho \in \{2, 4\}$  for various bivariate and trivariate data-generating models and  $n \in \{10, 20, \dots, 100\}$ . Specifically, we considered:

- bivariate normal copulas with a Kendall's tau  $\tau \in \{-0.75, -0.5, -0.25, 0.25, 0.5, 0.75\}$ ,
- non-exchangeable bivariate Khoudraji–Clayton copulas with first shape parameter  $s_1 \in \{0.4, 0.6, 0.8\}$ , second shape parameter  $s_2 = 0.95$  and Clayton copula parameter equal to 6 (see Section 3.4.2 in Hofert et al. (2018) and Fig. 3.19 therein in particular),
- trivariate Gumbel–Hougaard copulas whose bivariate margins have a Kendall's tau  $\tau \in \{0.25, 0.5, 0.75\}$ ,
- the trivariate  $t$  copula with 4 degrees of freedom and pairwise correlation parameters  $\rho_{12} = -0.2$ ,  $\rho_{13} = 0.5$  and  $\rho_{23} = 0.4$  as in Segers et al. (2017),
- the trivariate nested Archimedean copula with Frank generators and with a Kendall's tau equal to 0.3 at the upper node and 0.6 at the lower node as in Segers et al. (2017),
- the trivariate nested Archimedean copula with Frank generators and with a Kendall's tau equal to 0.5 at the upper node and 0.8 at the lower node.

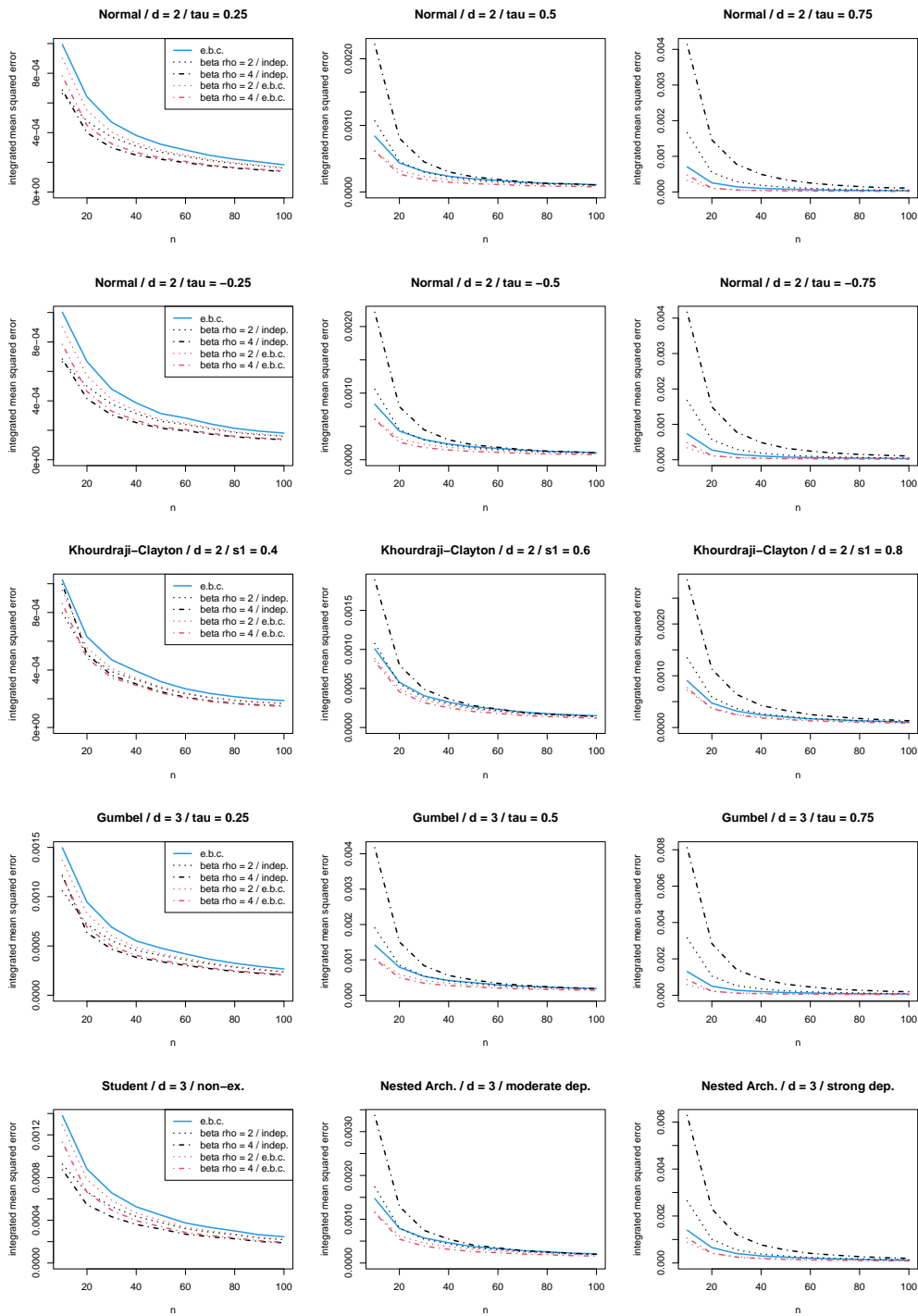


FIGURE 3.4: Estimated integrated mean squared errors against  $n \in \{10, 20, \dots, 100\}$  of five estimators of  $C$  for the bivariate and trivariate data-generating models considered in Section 3.5.2.

The estimated integrated mean squared errors are plotted against  $n \in \{10, 20, \dots, 100\}$  in the graphs of Fig. 3.4 for the estimators (other than  $C_{1:n}^{\text{beta}}$ ) that have beta smoothing survival margins. As already mentioned, the estimated integrated mean squared errors of the estimators having scaled beta-binomial smoothing survival margins (all other settings being equal) are



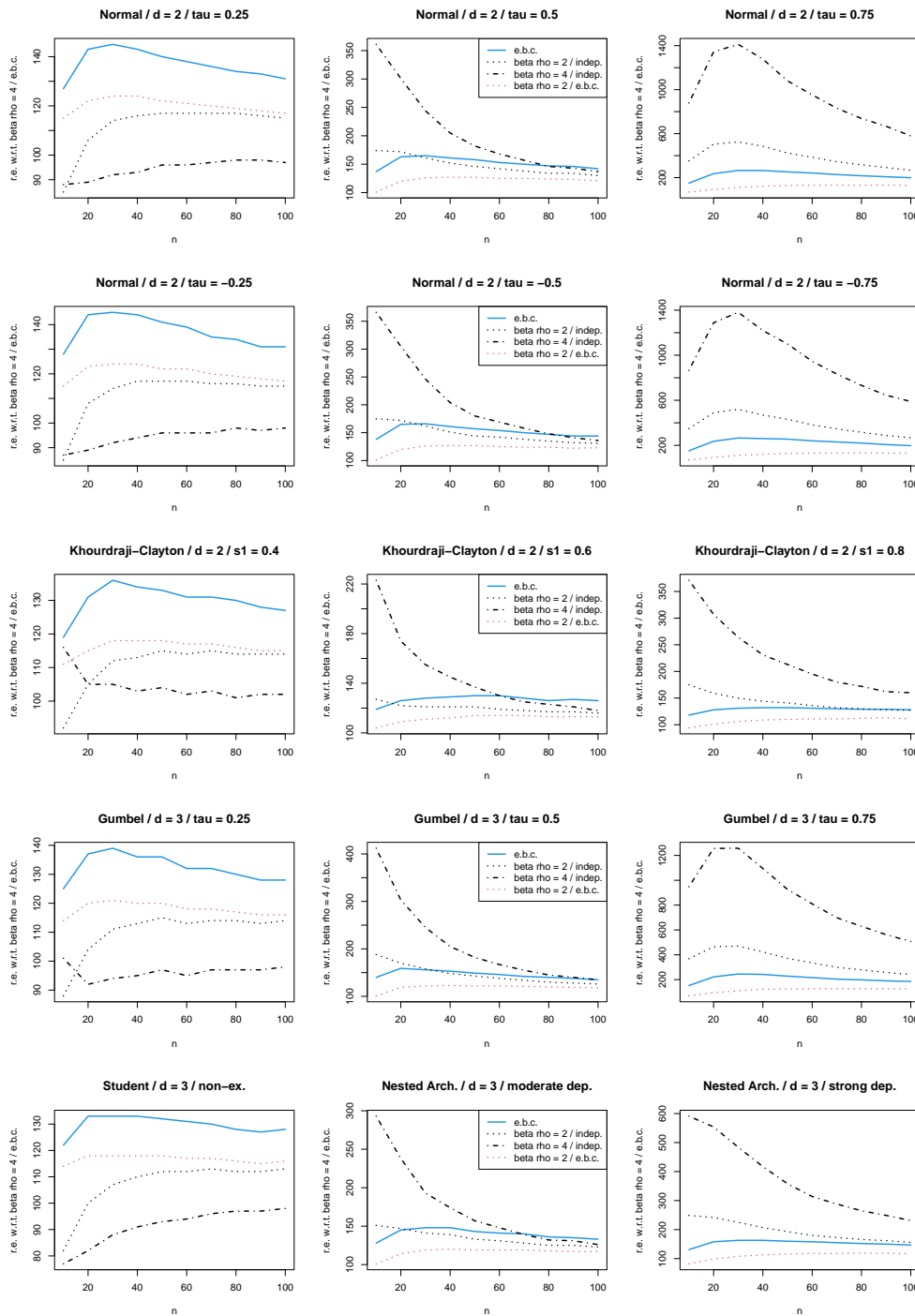


FIGURE 3.5: Estimated relative efficiencies (r.e.) with respect to the estimator  $C_{1:n}^{\bar{\beta}, C_{1:n}^{\beta}}$  with  $\rho = 4$  against  $n \in \{10, 20, \dots, 100\}$  for the four other estimators of  $C$  appearing in Fig. 3.4 for the bivariate and trivariate data-generating models considered in Section 3.5.2.

almost identical. Again, as can be seen from Fig. 3.4, the estimators  $C_{1:n}^{\bar{\beta}, C_{1:n}^{\beta}}$  and  $C_{1:n}^{\bar{\beta}, C_{1:n}^{\beta}}$  with  $\rho \in \{2, 4\}$  are uniformly better than the empirical beta copula in terms integrated mean squared error for all the data generating models under consideration, the best of the two being overall the one with  $\rho = 4$ , except for very small  $n$ . For an easier evaluation of the possible gain resulting

from the use of the latter estimator, Fig. 3.5 reports estimated relative efficiencies with respect to the estimator  $C_{1:n}^{\bar{\beta}, C_{1:n}^{\text{beta}}}$  with  $\rho = 4$  against  $n \in \{10, 20, \dots, 100\}$  (the relative efficiency of an estimator with respect to  $C_{1:n}^{\bar{\beta}, C_{1:n}^{\text{beta}}}$  with  $\rho = 4$  is simply the ratio of its integrated mean squared error to that of  $C_{1:n}^{\bar{\beta}, C_{1:n}^{\text{beta}}}$  with  $\rho = 4$  multiplied by 100). As already observed in the previous set of experiments, the estimators  $C_{1:n}^{\bar{\beta}, \Pi}$  and  $C_{1:n}^{\beta, \Pi}$  with  $\rho \in \{2, 4\}$  can be of interest only when the true copula is close to the independence copula  $\Pi$ .

### 3.6 Asymptotics of related sequential empirical processes

Recall the definition given in (3.5) of the smooth, possibly data-adaptive empirical copulas  $C_{1:n}^\nu$  investigated in this work. Conditions under which empirical processes of the form  $\sqrt{n}(C_{1:n}^\nu - C)$  converge weakly will be a consequence of the sequential results to be stated in this section. The reason for considering this more general setting follows from our intention to apply the studied smooth estimators in change-point analysis in a forthcoming project.

Before providing the asymptotic theory for the related sequential empirical processes, we start by briefly explaining how the quantities introduced in Section 3.2 can be defined from substretches of the available stretch of observations  $\mathbf{X}_{1:n} = (\mathbf{X}_1, \dots, \mathbf{X}_n)$ . Given a substretch  $\mathbf{X}_{k:l} = (\mathbf{X}_k, \dots, \mathbf{X}_l)$ ,  $1 \leq k \leq l \leq n$ , let  $F_{k:l,j}$  be the empirical d.f. computed from the  $j$ th component subsample  $X_{kj}, \dots, X_{lj}$ ,  $j \in \{1, \dots, d\}$  of  $\mathbf{X}_{k:l}$ . Then, let  $R_{ij}^{k:l} = (l - k + 1)F_{k:l,j}(X_{ij})$  be the (maximal) rank of  $X_{ij}$  among  $X_{kj}, \dots, X_{lj}$ . Furthermore, let  $\mathbf{R}_i^{k:l} = (R_{i1}^{k:l}, \dots, R_{id}^{k:l})$  and  $\hat{U}_i^{k:l} = \mathbf{R}_i^{k:l} / (l - k + 1)$ ,  $i \in \{k, \dots, l\}$ , be the multivariate ranks and the multivariate scaled ranks, respectively, obtained from  $\mathbf{X}_{k:l}$ . The empirical copula  $C_{k:l}$  of  $\mathbf{X}_{k:l}$  at  $\mathbf{u} = (u_1, \dots, u_d) \in [0, 1]^d$  is then

$$C_{k:l}(\mathbf{u}) = \frac{1}{l - k + 1} \sum_{i=k}^l \prod_{j=1}^d \mathbf{1} \left( \frac{R_{ij}^{k:l}}{l - k + 1} \leq u_j \right) = \frac{1}{l - k + 1} \sum_{i=k}^l \mathbf{1}(\hat{U}_i^{k:l} \leq \mathbf{u}), \quad (3.28)$$

and, by analogy with (3.5), the corresponding smooth version  $C_{k:l}^\nu$  based on the possibly random smoothing distributions  $\nu_{\mathbf{u}}^{\mathbf{X}_{k:l}}$ ,  $\mathbf{u} \in [0, 1]^d$ , is defined by

$$C_{k:l}^\nu(\mathbf{u}) = \int_{[0,1]^d} C_{k:l}(\mathbf{w}) d\nu_{\mathbf{u}}^{\mathbf{X}_{k:l}}(\mathbf{w}), \quad \mathbf{u} \in [0, 1]^d. \quad (3.29)$$

Note that, if  $k > l$ , we adopt the convention that  $C_{k:l} = C_{k:l}^\nu = 0$  and that, for any  $\mathbf{u} \in [0, 1]^d$ ,  $\nu_{\mathbf{u}}^{\mathbf{X}_{k:l}}$  is the Dirac measure at  $\mathbf{u}$ .

We can now define the corresponding sequential empirical copula processes which are the main focus of this section. Let  $\Delta = \{(s, t) \in [0, 1]^2 : s \leq t\}$  and let  $\lambda_n(s, t) = ([nt] - [ns])/n$ ,  $(s, t) \in \Delta$ . The two-sided sequential empirical copula process was defined in Bücher and Kojadinovic (2016) by

$$\mathbb{C}_n(s, t, \mathbf{u}) = \sqrt{n} \lambda_n(s, t) \{C_{[ns]+1:[nt]}(\mathbf{u}) - C(\mathbf{u})\}, \quad (s, t) \in \Delta, \mathbf{u} \in [0, 1]^d, \quad (3.30)$$

where  $C_{[ns]+1:[nt]}$  is generically defined in (3.28). The two-sided sequential empirical copula process corresponding to the smooth empirical copula in (3.29) is then naturally defined by

$$\mathbb{C}_n^\nu(s, t, \mathbf{u}) = \sqrt{n}\lambda_n(s, t)\{C_{[ns]+1:[nt]}^\nu(\mathbf{u}) - C(\mathbf{u})\}, \quad (s, t) \in \Delta, \mathbf{u} \in [0, 1]^d. \quad (3.31)$$

The aim of this section is to establish the asymptotics of the latter.

As frequently done in the literature, we consider the following nonrestrictive condition on the unknown copula  $C$  initially proposed in Segers (2012).

**Condition 3.6.1** (Smooth partial derivatives). For any  $j \in \{1, \dots, d\}$ , the partial derivative  $\dot{C}_j = \partial C / \partial u_j$  exists and is continuous on the set  $\{\mathbf{u} \in [0, 1]^d : u_j \in (0, 1)\}$ .

The latter condition can be considered nonrestrictive because, as shall become clear in the forthcoming developments, it is necessary for the candidate weak limit of both  $\mathbb{C}_n$  and  $\mathbb{C}_n^\nu$  to exist pointwise and have continuous trajectories. In the rest of the paper, following Bücher and Volgushev (2013), for any  $j \in \{1, \dots, d\}$ ,  $\dot{C}_j$  is arbitrarily defined to be zero on the set  $\{\mathbf{u} \in [0, 1]^d : u_j \in \{0, 1\}\}$ , which implies that, under Condition 3.6.1,  $\dot{C}_j$  is defined on the whole of  $[0, 1]^d$ .

In view of the mathematical derivations carried out in Segers et al. (2017) to study the asymptotics of the empirical beta copula process and with (3.4) in mind, it seems particularly meaningful to impose the following condition on the smoothing distributions involved in the definition of the smooth empirical copulas under consideration.

**Condition 3.6.2** (Variance condition). There exists a constant  $\kappa > 0$  such that, for any  $n \in \mathbb{N}$ ,  $\mathbf{x} \in (\mathbb{R}^d)^n$ ,  $\mathbf{u} \in [0, 1]^d$  and  $j \in \{1, \dots, d\}$ ,  $\text{Var}(W_{j, u_j}^{\mathbf{x}}) \leq \kappa u_j(1 - u_j)/n$ .

*Remark 3.6.3.* As verified at the beginning of Section 3.4.1, the previous condition is clearly satisfied if, for any  $n \in \mathbb{N}$ ,  $\mathbf{x} \in (\mathbb{R}^d)^n$ ,  $\mathbf{u} \in [0, 1]^d$  and  $j \in \{1, \dots, d\}$ ,  $W_{j, u_j}^{\mathbf{x}} = S_{n, j, u_j}/n$ , where  $S_{n, j, u_j}$  is Binomial( $n, u_j$ ). From (3.16) and (3.21), we see that it will also be satisfied for smoothing distributions with scaled beta-binomial and beta survival margins as defined in Sections 3.4.1 and 3.4.2 provided that the dispersion parameter  $\rho$  is constant. For versions of these models in which  $\rho$  is data-adaptive, one needs to assume that there exists a constant  $\gamma > 0$ , such for any  $n \in \mathbb{N}$  and  $\mathbf{x} \in (\mathbb{R}^d)^n$ , the dispersion parameter  $\rho = \rho^{\mathbf{x}}$  (which may thus depend on  $\mathbf{x}$ ) is bounded by  $\gamma$ .

In the rest of this section, unless stated otherwise, all convergences are with respect to  $n \rightarrow \infty$ . Furthermore, the arrow ‘ $\rightsquigarrow$ ’ denotes weak convergence in the sense of Definition 1.3.3 in van der Vaart and Wellner (2000) and, given a set  $T$ ,  $\ell^\infty(T)$  (resp.  $\mathcal{C}(T)$ ) represents the space of all bounded (resp. continuous) real-valued functions on  $T$  equipped with the uniform metric. The following result is proven in Appendix 3.8.3.

**Theorem 3.6.4** (Asymptotics of  $\mathbb{C}_n^\nu$ ). Assume that Conditions 3.6.1 and 3.6.2 hold, and that  $\mathbb{C}_n \rightsquigarrow \mathbb{C}_C$  in  $\ell^\infty(\Delta \times [0, 1]^d)$ , where the trajectories of the limiting process  $\mathbb{C}_C$  are continuous

almost surely. Then,

$$\sup_{\substack{(s,t) \in \Delta \\ \mathbf{u} \in [0,1]^d}} |\mathbb{C}_n^\nu(s, t, \mathbf{u}) - \mathbb{C}_n(s, t, \mathbf{u})| = o_{\mathbb{P}}(1).$$

Consequently,  $\mathbb{C}_n^\nu \rightsquigarrow \mathbb{C}_C$  in  $\ell^\infty(\Delta \times [0, 1]^d)$ .

Roughly speaking, the smooth sequential empirical copula process  $\mathbb{C}_n^\nu$  in (3.31) is asymptotically equivalent to the classical sequential empirical copula process  $\mathbb{C}_n$  in (3.30) when the latter converges weakly to a limiting process whose trajectories are continuous almost surely. From Section 3 in Bücher and Kojadinovic (2016), a key assumption for such a convergence to hold concerns the weak convergence of the two-sided sequential empirical process

$$\mathbb{B}_n(s, t, \mathbf{u}) = \frac{1}{\sqrt{n}} \sum_{i=[ns]+1}^{\lfloor nt \rfloor} \{\mathbf{1}(\mathbf{U}_i \leq \mathbf{u}) - C(\mathbf{u})\}, \quad (s, t, \mathbf{u}) \in \Delta \times [0, 1]^d, \quad (3.32)$$

where  $\mathbf{U}_1, \dots, \mathbf{U}_n$  is the unobservable sample obtained from  $\mathbf{X}_{1:n} = (\mathbf{X}_1, \dots, \mathbf{X}_n)$  by the probability integral transformations  $U_{ij} = F_j(X_{ij})$ ,  $i \in \{1, \dots, n\}$ ,  $j \in \{1, \dots, d\}$ , and with the convention that  $\mathbb{B}_n(s, t, \cdot) = 0$  if  $\lfloor nt \rfloor - \lfloor ns \rfloor = 0$ . Specifically, the following condition was considered in Bücher and Kojadinovic (2016).

**Condition 3.6.5** (Weak convergence of  $\mathbb{B}_n(0, \cdot, \cdot)$ ). The sequential empirical process  $\mathbb{B}_n(0, \cdot, \cdot)$  converges weakly in  $\ell^\infty([0, 1]^{d+1})$  to a tight centered Gaussian process  $\mathbb{Z}_C$  concentrated on

$$\{f \in \mathcal{C}([0, 1]^{d+1}) : f(s, \mathbf{u}) = 0 \text{ if one of the components of } (s, \mathbf{u}) \text{ is } 0, \text{ and } f(s, 1, \dots, 1) = 0 \text{ for all } s \in (0, 1]\}.$$

Under Condition 3.6.5, it immediately follows from the continuous mapping theorem that  $\mathbb{B}_n \rightsquigarrow \mathbb{B}_C$  in  $\ell^\infty(\Delta \times [0, 1]^d)$ , where

$$\mathbb{B}_C(s, t, \mathbf{u}) = \mathbb{Z}_C(t, \mathbf{u}) - \mathbb{Z}_C(s, \mathbf{u}), \quad (s, t, \mathbf{u}) \in \Delta \times [0, 1]^d. \quad (3.33)$$

When the underlying stationary time series  $(\mathbf{X}_i)_{i \in \mathbb{Z}}$  consists of independent random vectors, Condition 3.6.5 is a direct consequence, for instance, of Theorem 2.12.1 in van der Vaart and Wellner (2000). More generally, Condition 3.6.5 also holds when  $(\mathbf{X}_i)_{i \in \mathbb{Z}}$  is strongly mixing. Given a stationary time series  $(\mathbf{Y}_i)_{i \in \mathbb{Z}}$ , denote by  $\mathcal{F}_j^k$  the  $\sigma$ -field generated by  $(\mathbf{Y}_i)_{j \leq i \leq k}$ ,  $j, k \in \mathbb{Z} \cup \{-\infty, +\infty\}$ , and recall that the strong mixing coefficients corresponding to the stationary sequence  $(\mathbf{Y}_i)_{i \in \mathbb{Z}}$  are then defined by

$$\alpha_r^{\mathbf{Y}} = \sup_{A \in \mathcal{F}_{-\infty}^0, B \in \mathcal{F}_r^{+\infty}} |\mathbb{P}(A \cap B) - \mathbb{P}(A)\mathbb{P}(B)|, \quad r \in \mathbb{N}, r > 0,$$

with  $\alpha_0^{\mathbf{Y}} = 1/2$  and that the sequence  $(\mathbf{Y}_i)_{i \in \mathbb{Z}}$  is said to be strongly mixing if  $\alpha_r^{\mathbf{Y}} \rightarrow 0$  as  $r \rightarrow \infty$ . Then, if the stretch  $\mathbf{U}_1, \dots, \mathbf{U}_n$  is drawn from a time series  $(\mathbf{U}_i)_{i \in \mathbb{Z}}$  whose strong mixing

coefficients satisfy  $\alpha_r^U = O(r^{-a})$  with  $a > 1$  (which occurs if the strong mixing coefficients of the time series  $(\mathbf{X}_i)_{i \in \mathbb{Z}}$  satisfy  $\alpha_r^{\mathbf{X}} = O(r^{-a})$  with  $a > 1$ ), Theorem 1 in Bücher (2015) implies that Condition 3.6.5 holds with the covariance function of the process  $\mathbb{Z}_C$  being

$$\text{Cov}\{\mathbb{Z}_C(s, \mathbf{u}), \mathbb{Z}_C(t, \mathbf{v})\} = \min(s, t) \sum_{k \in \mathbb{Z}} \text{Cov}\{\mathbf{1}(U_0 \leq \mathbf{u}), \mathbf{1}(U_k \leq \mathbf{v})\}.$$

Recall that, for any  $j \in \{1, \dots, d\}$  and any  $\mathbf{u} \in [0, 1]^d$ ,  $\mathbf{u}^{(j)}$  is the vector of  $[0, 1]^d$  defined by  $u_i^{(j)} = u_j$  if  $i = j$  and 1 otherwise. The following result is then an immediate consequence of Theorem 3.4 in Bücher and Kojadinovic (2016) and Proposition 3.3 of Bücher et al. (2014), and can be used to obtain a straightforward corollary of Theorem 3.6.4.

**Theorem 3.6.6** (Asymptotics of  $\mathbb{C}_n$ ). *Under Conditions 3.6.1 and 3.6.5,*

$$\sup_{(s, t, \mathbf{u}) \in \Delta \times [0, 1]^d} \left| \mathbb{C}_n(s, t, \mathbf{u}) - \tilde{\mathbb{C}}_n(s, t, \mathbf{u}) \right| = o_{\mathbb{P}}(1),$$

where

$$\tilde{\mathbb{C}}_n(s, t, \mathbf{u}) = \mathbb{B}_n(s, t, \mathbf{u}) - \sum_{j=1}^d \dot{C}_j(\mathbf{u}) \mathbb{B}_n(s, t, \mathbf{u}^{(j)}), \quad (s, t, \mathbf{u}) \in \Delta \times [0, 1]^d,$$

and  $\mathbb{B}_n$  is defined in (3.32). Consequently,  $\mathbb{C}_n \rightsquigarrow \mathbb{C}_C$  in  $\ell^\infty(\Delta \times [0, 1]^d)$ , where

$$\mathbb{C}_C(s, t, \mathbf{u}) = \mathbb{B}_C(s, t, \mathbf{u}) - \sum_{j=1}^d \dot{C}_j(\mathbf{u}) \mathbb{B}_C(s, t, \mathbf{u}^{(j)}), \quad (s, t, \mathbf{u}) \in \Delta \times [0, 1]^d,$$

and  $\mathbb{B}_C$  is defined in (3.33).

### 3.7 Concluding remarks

A broad class of smooth, possibly data-adaptive empirical copulas containing the empirical beta copula was studied. A good general choice within this class appears to be the estimator  $C_{1:n}^{\beta, C_{1:n}^{\text{beta}}}$  generically defined in (3.18) that uses the empirical beta copula as survival copula of the smoothing distribution and scaled beta-binomial smoothing survival margins with dispersion parameter  $\rho = 4$ . From Corollary 3.3.8, the latter is a genuine copula in the absence of ties in the component samples. In our bivariate and trivariate Monte Carlo experiments, it was found to be uniformly better than the empirical beta copula for all the considered data-generating models. However, its better finite-sample performance compared to the empirical beta copula comes at the price of a higher computational cost since this more complex estimator uses the empirical beta copula as a pilot estimator of the true unknown copula. Specifically, we see from (3.18) that one estimation of the unknown copula at a given point  $\mathbf{u} \in [0, 1]^d$  will require  $n$  evaluations of the empirical beta copula. Whether this overhead is acceptable for a given

application can be assessed using the R implementation of the estimator available on the web page of the first author.

In addition to these finite-sample results, conditions under which sequential empirical copula processes constructed from the studied general class of smooth, possibly data-adaptive estimators converge weakly were provided. In a forthcoming project, these results will be used to show the asymptotic validity of a sequential dependent multiplier bootstrap [Bücher and Kojadinovic \(2016\)](#); [Bücher et al. \(2014\)](#) allowing to apply the proposed smooth, possibly data-adaptive empirical copulas to change-point detection in a time series setting.

## 3.8 Proofs

### 3.8.1 Proofs of Lemma 3.2.1, Proposition 3.3.2, Proposition 3.3.7 and Proposition 3.3.10

**Proof of Lemma 3.2.1.** For any  $\mathbf{u} \in [0, 1]^d$ ,  $n \in \mathbb{N}$  and  $r_1, \dots, r_d \in \{1, \dots, n\}$ , we have

$$\begin{aligned} \int_{[0,1]^d} \prod_{j=1}^d \mathbf{1}(r_j/n \leq w_j) d\mu_{n,\mathbf{u}}(\mathbf{w}) &= \mathbb{E} \left\{ \prod_{j=1}^d \mathbf{1}(r_j/n \leq S_{n,j,u_j}/n) \right\} \\ &= \prod_{j=1}^d \mathbb{E} \{ \mathbf{1}(r_j \leq S_{n,j,u_j}) \} = \prod_{j=1}^d \mathbb{P}(S_{n,j,u_j} \geq r_j) \\ &= \prod_{j=1}^d \{ 1 - \mathbb{P}(S_{n,j,u_j} \leq r_j - 1) \} = \prod_{j=1}^d \left\{ 1 - \sum_{s=0}^{r_j-1} \binom{n}{s} u_j^s (1-u_j)^{n-s} \right\} \\ &= \prod_{j=1}^d \left\{ \sum_{s=0}^n \binom{n}{s} u_j^s (1-u_j)^{n-s} - \sum_{s=0}^{r_j-1} \binom{n}{s} u_j^s (1-u_j)^{n-s} \right\} \\ &= \prod_{j=1}^d \sum_{s=r_j}^n \binom{n}{s} u_j^s (1-u_j)^{n-s} = \prod_{j=1}^d \text{beta}_{n,r_j}(u_j). \end{aligned}$$

■

**Proof of Proposition 3.3.2.** Fix  $j \in \{1, \dots, d\}$  and  $\mathbf{u} \in [0, 1]^d$ . Then, from (3.4), all the components of  $\mathbf{W}_{\mathbf{u}^{(j)}}^{\mathbf{x}^{1:n}}$  except possibly the  $j$ th are deterministic and equal to 1. It follows

from (3.6) and (3.7) that, almost surely,

$$\begin{aligned} C_{1:n}^\nu(\mathbf{u}^{(j)}) &= \frac{1}{n} \sum_{i=1}^n \mathcal{K}_{\mathbf{R}_i^{1:n}}^{\mathbf{x}^{1:n}}(\mathbf{u}^{(j)}) = \frac{1}{n} \sum_{i=1}^n \int_{[0,1]^d} \mathbf{1}(R_{ij}^{1:n}/n \leq w_j) d\nu_{\mathbf{u}^{1:n}}^{\mathbf{x}^{1:n}}(\mathbf{w}) \\ &= \frac{1}{n} \sum_{i=1}^n \int_{[0,1]^d} \mathbf{1}(i/n \leq w_j) d\nu_{\mathbf{u}^{1:n}}^{\mathbf{x}^{1:n}}(\mathbf{w}) = \frac{1}{n} \sum_{i=1}^n \mathbb{E} \left\{ \mathbf{1} \left( \frac{i}{n} \leq W_{j,u_j}^{\mathbf{x}^{1:n}} \right) \mid \mathbf{x}^{1:n} \right\} \\ &= \mathbb{E} \left\{ \frac{1}{n} \sum_{i=1}^n \mathbf{1} \left( i \leq nW_{j,u_j}^{\mathbf{x}^{1:n}} \right) \mid \mathbf{x}^{1:n} \right\} = \mathbb{E} \left( \frac{\lfloor nW_{j,u_j}^{\mathbf{x}^{1:n}} \rfloor}{n} \mid \mathbf{x}^{1:n} \right). \end{aligned}$$

■

**Proofs of Propositions 3.3.7 and 3.3.10.** We only provide the proof of Proposition 3.3.7 as the proof of Proposition 3.3.10 is very similar.

Let us check that  $C_{1:n}^\nu$  in (3.10) satisfies the four properties listed in Theorem 1.2.11 of Durante and Sempi (2015) necessary for it to be a multivariate d.f. From (3.9), under Condition 3.3.5, for any  $\mathbf{u} \in [0,1]^d$  and  $i \in \{1, \dots, n\}$ , one has that, almost surely

$$\mathcal{K}_{\mathbf{R}_i^{1:n}}^{\mathbf{x}^{1:n}}(\mathbf{u}) = \bar{C}^{\mathbf{x}^{1:n}}[\bar{\mathcal{F}}_{1,u_1}^{\mathbf{x}^{1:n}}\{(R_{i1}^{1:n} - 1)/n\}, \dots, \bar{\mathcal{F}}_{d,u_d}^{\mathbf{x}^{1:n}}\{(R_{id}^{1:n} - 1)/n\}].$$

Since, for any  $\mathbf{u} \in [0,1]^d$  and  $j \in \{1, \dots, d\}$ ,  $W_{j,u_j}^{\mathbf{x}^{1:n}} = 0$  if the  $j$ th coordinate of  $\mathbf{u}$  is zero and  $\bar{C}^{\mathbf{x}^{1:n}}$  is a copula, we obtain that, with probability 1,  $\mathcal{K}_{\mathbf{R}_i^{1:n}}^{\mathbf{x}^{1:n}}(\mathbf{u}) = 0$  for all  $i \in \{1, \dots, n\}$  if at least one coordinate of  $\mathbf{u}$  is zero. The latter implies that the estimator  $C_{1:n}^\nu$  in (3.10) is grounded. Similarly, if  $\mathbf{u} = \mathbf{1} \in [0,1]^d$ , with probability 1  $W_{j,u_j}^{\mathbf{x}^{1:n}} = 1$  for all  $j \in \{1, \dots, d\}$  and  $\mathcal{K}_{\mathbf{R}_i^{1:n}}^{\mathbf{x}^{1:n}}(\mathbf{u}) = 1$  for all  $i \in \{1, \dots, n\}$ , which implies that  $C_{1:n}^\nu(\mathbf{1}) = 1$  almost surely.

From Condition 3.3.6 and the fact that  $\bar{C}^{\mathbf{x}}$  is a copula for all  $\mathbf{x} \in (\mathbb{R}^d)^n$ , we have that, for any  $\mathbf{x} \in (\mathbb{R}^d)^n$ ,  $j \in \{1, \dots, d\}$ ,  $\mathbf{u} \in [0,1]^d$  and  $\mathbf{w} \in [0,1]^d$ , the function from  $[0,1]$  to  $[0,1]$  defined by

$$t \mapsto \bar{C}^{\mathbf{x}}\{\bar{\mathcal{F}}_{1,u_1}^{\mathbf{x}}(w_1), \dots, \bar{\mathcal{F}}_{j,u_{j-1}}^{\mathbf{x}}(w_{j-1}), \bar{\mathcal{F}}_{j,t}^{\mathbf{x}}(w_j), \bar{\mathcal{F}}_{j,u_{j+1}}^{\mathbf{x}}(w_{j+1}), \dots, \bar{\mathcal{F}}_{d,u_d}^{\mathbf{x}}(w_d)\}$$

is right-continuous. The latter implies that, with probability 1, for any  $j \in \{1, \dots, d\}$  and  $\mathbf{u} \in [0,1]^d$ , the function from  $[0,1]$  to  $[0,1]$  defined by  $t \mapsto C_{1:n}^\nu(u_1, \dots, u_{j-1}, t, u_{j+1}, \dots, u_d)$  is right-continuous as well.

For any  $\mathbf{a}, \mathbf{b} \in [0,1]^d$  such that  $\mathbf{a} \leq \mathbf{b}$ , let  $(\mathbf{a}, \mathbf{b}) = (a_1, b_1) \times \dots \times (a_d, b_d)$  and, for any function  $H$  on  $[0,1]^d$ , let the  $H$ -volume of  $(\mathbf{a}, \mathbf{b})$  be defined by

$$\Delta_{(\mathbf{a}, \mathbf{b})} H = \sum_{\mathbf{i} \in \{0,1\}^d} (-1)^{\sum_{j=1}^d i_j} H(a_1^{i_1} b_1^{1-i_1}, \dots, a_d^{i_d} b_d^{1-i_d}).$$

Then, some thought reveals that, for any  $\mathbf{a}, \mathbf{b} \in [0,1]^d$  such that  $\mathbf{a} \leq \mathbf{b}$ , Condition 3.3.6 and the fact that  $\bar{C}^{\mathbf{x}}$  is  $d$ -increasing (see, e.g., Durante and Sempi, 2015, Definition 1.2.9) for all  $\mathbf{x} \in (\mathbb{R}^d)^n$  imply that, with probability 1,  $\Delta_{(\mathbf{a}, \mathbf{b})} \mathcal{K}_{\mathbf{R}_i^{1:n}}^{\mathbf{x}^{1:n}} \geq 0$  for all  $i \in \{1, \dots, n\}$ , which in turn

implies that  $\Delta_{(a,b]}C_{1:n}^\nu = \frac{1}{n} \sum_{i=1}^n \Delta_{(a,b]} \mathcal{K}_{\mathbf{R}_i^{1:n}}^{\mathcal{X}_{1:n}} \geq 0$  almost surely, and thus that the estimator  $C_{1:n}^\nu$  is  $d$ -increasing with probability 1.

The desired conclusion finally follows from Theorem 1.2.11 in [Durante and Sempi \(2015\)](#). ■

### 3.8.2 Proofs of Propositions 3.4.1, 3.4.2 and 3.4.4

The proofs of Propositions 3.4.1, 3.4.2 and 3.4.4 are based on six lemmas which we prove first.

**Lemma 3.8.1.** *For any  $n \in \mathbb{N}$  and  $w \in [0, n)$ , the function  $t \mapsto \bar{B}_{n,t}(w)$  is continuous on  $[0, 1]$ .*

**Proof.** Fix  $n \in \mathbb{N}$  and  $w \in [0, n)$ . The desired result is an immediate consequence of the fact that

$$\bar{B}_{n,t}(w) = 1 - \sum_{k=0}^{\lfloor w \rfloor} \binom{n}{k} t^k (1-t)^{n-k}, \quad t \in [0, 1],$$

is a polynomial in  $t$ . ■

**Lemma 3.8.2.** *For any  $n \in \mathbb{N}$ ,  $\rho \in (0, n)$  and  $w \in (0, 1)$ , the function  $t \mapsto \bar{\beta}_{n,t,\rho}(w)$  is continuous on  $[0, 1]$ .*

**Proof.** Fix  $n \in \mathbb{N}$ ,  $\rho \in (0, n)$  and  $w \in (0, 1)$  and let  $\tau = (n - \rho)/\rho > 0$ . From the properties of the beta distribution  $\text{Beta}(t\tau, (1-t)\tau)$ , for any  $t \in (0, 1)$ ,

$$\bar{\beta}_{n,t,\rho}(w) = 1 - \int_0^w f_{t,\tau}(x) dx,$$

where

$$f_{t,\tau}(x) = \frac{\Gamma(\tau)}{\Gamma(t\tau)\Gamma\{(1-t)\tau\}} x^{t\tau-1}(1-x)^{(1-t)\tau-1}, \quad x \in (0, 1),$$

and  $\Gamma$  is the gamma function. Fix  $t_0 \in (0, 1)$  and let us verify that  $t \mapsto \bar{\beta}_{n,t,\rho}(w)$  is continuous at  $t_0$ . To do so, let  $t_m$  be an arbitrary sequence in  $(0, 1)$  such that  $t_m \rightarrow t_0$  as  $m \rightarrow \infty$ . By continuity of  $\Gamma$  on  $(0, \infty)$ , we have that, for any  $x \in (0, 1)$ ,  $f_{t_m,\tau}(x) \rightarrow f_{t_0,\tau}(x)$  as  $m \rightarrow \infty$ . Furthermore, since there exists a constant  $K > 0$  such that  $\Gamma(z) \geq K$  for all  $z \in (0, \infty)$ , we have that, for any  $m \in \mathbb{N}$  and  $x \in (0, 1)$ ,  $f_{t_m,\tau}(x) \leq \Gamma(\tau)/K^2$ . The dominated convergence theorem then implies that  $\bar{\beta}_{n,t_m,\rho}(w) \rightarrow \bar{\beta}_{n,t_0,\rho}(w)$  as  $m \rightarrow \infty$ , which implies that the function  $t \mapsto \bar{\beta}_{n,t,\rho}(w)$  is continuous on  $(0, 1)$ . Furthermore, from Markov's inequality, we have that, for any  $t \in (0, 1)$ ,  $\bar{\beta}_{n,t,\rho}(w) \leq t/w$ , which combined with the fact that  $\bar{\beta}_{n,0,\rho}(w) = 0$ , implies that  $\lim_{t \rightarrow 0^+} \bar{\beta}_{n,t,\rho}(w) = \bar{\beta}_{n,0,\rho}(w)$ , and therefore that the function  $t \mapsto \bar{\beta}_{n,t,\rho}(w)$  is continuous at 0. Finally, using symmetry properties of the density of the  $\text{Beta}(t\tau, (1-t)\tau)$ , some thought reveals that, for any  $t \in (0, 1)$ ,  $\bar{\beta}_{n,1-t,\rho}(w) = \beta_{n,t,\rho}(1-w) = 1 - \bar{\beta}_{n,t,\rho}(1-w)$ , which implies that  $\lim_{t \rightarrow 1^-} \bar{\beta}_{n,t,\rho}(w) = \lim_{t \rightarrow 0^+} \bar{\beta}_{n,1-t,\rho}(w) = 1 - \lim_{t \rightarrow 0^+} \bar{\beta}_{n,t,\rho}(1-w) = 1 - \bar{\beta}_{n,1,\rho}(w)$  since, from Markov's inequality, for any  $t \in (0, 1)$ ,  $\bar{\beta}_{n,t,\rho}(1-w) \leq t/(1-w)$ . ■



**Lemma 3.8.3.** *For any  $n \in \mathbb{N}$ ,  $\rho \in (0, n)$  and  $w \in [0, n)$ , the function  $t \mapsto \bar{\mathcal{B}}_{n,t,\rho}(w)$  is continuous on  $[0, 1]$ .*

**Proof.** Fix  $n \in \mathbb{N}$ ,  $\rho \in (1, n)$  and  $w \in [0, n)$  and let  $\tau = (n - \rho)/(\rho - 1) > 0$ . From the properties of the beta-binomial distribution  $\text{Beta-Binomial}(n, t\tau, (1 - t)\tau)$ , for any  $t \in (0, 1)$ , we have that

$$\bar{\mathcal{B}}_{n,t,\rho}(w) = 1 - \sum_{k=0}^{\lfloor w \rfloor} \binom{n}{k} \frac{\Gamma(k + t\tau)\Gamma\{n - k + (1 - t)\tau\}\Gamma(\tau)}{\Gamma(n + \tau)\Gamma(t\tau)\Gamma\{(1 - t)\tau\}}.$$

Since  $\Gamma$  is continuous on  $(0, \infty)$ , we immediately obtain that  $t \mapsto \bar{\mathcal{B}}_{n,t,\rho}(w)$  is continuous on  $(0, 1)$ . For some  $t \in (0, 1)$ , let  $X$  be  $\text{Beta-Binomial}(n, t\tau, (1 - t)\tau)$ . Using Markov's inequality, we have that

$$\bar{\mathcal{B}}_{n,t,\rho}(w) = \mathbb{P}(X > w) = \mathbb{P}(X \geq \lfloor w \rfloor + 1) \leq \frac{nt}{\lfloor w \rfloor + 1}.$$

The latter display implies that  $\lim_{t \rightarrow 0^+} \bar{\mathcal{B}}_{n,t,\rho}(w) = 0 = \bar{\mathcal{B}}_{n,0,\rho}(w)$  since  $X$  is degenerate and equal to 0 if  $t = 0$ . It remains to verify that  $t \mapsto \bar{\mathcal{B}}_{n,t,\rho}(w)$  is continuous at 1. Using symmetry properties of the probability mass function of the  $\text{Beta-Binomial}(n, t\tau, (1 - t)\tau)$ , some thought reveals that, for any  $t \in (0, 1)$ ,  $\bar{\mathcal{B}}_{n,1-t,\rho}(w) = \mathcal{B}_{n,t,\rho}(n - w) = 1 - \bar{\mathcal{B}}_{n,t,\rho}(n - w)$ , which implies that  $\lim_{t \rightarrow 1^-} \bar{\mathcal{B}}_{n,t,\rho}(w) = \lim_{t \rightarrow 0^+} \bar{\mathcal{B}}_{n,1-t,\rho}(w) = 1 - \lim_{t \rightarrow 0^+} \bar{\mathcal{B}}_{n,t,\rho}(n - w) = 1 = \bar{\mathcal{B}}_{n,1,\rho}(w)$  since, from Markov's inequality, for any  $t \in (0, 1)$ ,  $\bar{\mathcal{B}}_{n,t,\rho}(n - w) \leq nt/(\lfloor n - w \rfloor + 1)$ . ■

**Lemma 3.8.4.** *For any  $n \in \mathbb{N}$  and  $w \in [0, n)$ , the function  $t \mapsto \bar{B}_{n,t}(w)$  is increasing on  $[0, 1]$ .*

**Proof.** Fix  $n \in \mathbb{N}$ . We need to prove that, for any  $w \in [0, n)$ ,  $\bar{B}_{n,t_1}(w) \leq \bar{B}_{n,t_2}(w)$  whenever  $0 \leq t_1 \leq t_2 \leq 1$ . Notice that, for any  $w \in [0, n)$ ,  $\bar{B}_{n,0}(w) = 0$  and  $\bar{B}_{n,1}(w) = 1$  so that it suffices to prove that, for any  $w \in [0, n)$ ,  $\bar{B}_{n,t_1}(w) \leq \bar{B}_{n,t_2}(w)$  whenever  $0 < t_1 \leq t_2 < 1$ . Fix  $0 < t_1 \leq t_2 < 1$  and let  $X$  be  $\text{Binomial}(n, t_1)$  and  $Y$  be  $\text{Binomial}(n, t_2)$ . Some thought reveals that the latter is proven if we show that  $X \leq_{\text{st}} Y$ , where  $\leq_{\text{st}}$  denotes the usual stochastic order. Indeed, by definition,  $X \leq_{\text{st}} Y$  if  $\mathbb{P}(X > x) \leq \mathbb{P}(Y > x)$  for all  $x \in \mathbb{R}$ . According for instance to Theorem 1.C.1 in [Shaked and Shanthikumar \(2007\)](#),  $X \leq_{\text{st}} Y$  will hold if  $X \leq_{\text{lr}} Y$ , where  $\leq_{\text{lr}}$  denotes the likelihood ratio order.

Let  $f$  be the probability mass function of  $X$  and let  $g$  be the probability mass function of  $Y$ . Then, by definition,

$$f(x) = \binom{n}{x} t_1^x (1 - t_1)^{n-x}, \quad g(x) = \binom{n}{x} t_2^x (1 - t_2)^{n-x}, \quad x \in \{0, \dots, n\}.$$

According for instance to Section 1.C.1 in [Shaked and Shanthikumar \(2007\)](#), to prove that  $X \leq_{\text{lr}} Y$ , by definition, we need to show that the function  $x \mapsto g(x)/f(x)$  is increasing on  $\{0, \dots, n\}$ . To prove the latter, it suffices to show that the function

$$h(x) = \frac{t_2^x (1 - t_2)^{n-x}}{t_1^x (1 - t_1)^{n-x}}, \quad x \in [0, n],$$

is increasing on  $[0, n]$ . To do so, we shall prove that the derivative of  $h$  is positive on  $[0, n]$ . Let  $l(x) = \log\{h(x)\}$ ,  $x \in [0, n]$ . Then, for any  $x \in [0, n]$ ,

$$\begin{aligned} \frac{d}{dx}\{h(x)\} &= \frac{d}{dx}\{e^{l(x)}\} = e^{l(x)} \frac{d}{dx}\{l(x)\} = e^{l(x)} \frac{d}{dx}[\log\{h(x)\}] \\ &= e^{l(x)} \frac{d}{dx} \left[ \log \left\{ \frac{t_2^x (1-t_2)^{n-x}}{t_1^x (1-t_1)^{n-x}} \right\} \right] \\ &= e^{l(x)} \frac{d}{dx} \{x \log(t_2) + (n-x) \log(1-t_2) - x \log(t_1) - (n-x) \log(1-t_1)\} \\ &= e^{l(x)} \log \left\{ \frac{t_2(1-t_1)}{t_1(1-t_2)} \right\} \geq 0, \end{aligned}$$

since  $0 < t_1 \leq t_2 < 1$  implies that  $t_2/t_1 \geq 1$  and that  $(1-t_1)/(1-t_2) \geq 1$ . ■

**Lemma 3.8.5.** *For any  $n \in \mathbb{N}$ ,  $\rho \in (0, n)$  and  $w \in (0, 1)$ , the function  $t \mapsto \bar{\beta}_{n,t,\rho}(w)$  is increasing on  $[0, 1]$ .*

**Proof.** Fix  $n \in \mathbb{N}$  and  $\rho \in (0, n)$ . We need to prove that, for any  $w \in (0, 1)$ ,  $\bar{\beta}_{n,t_1,\rho}(w) \leq \bar{\beta}_{n,t_2,\rho}(w)$  whenever  $0 \leq t_1 \leq t_2 \leq 1$ . Since, for any  $w \in (0, 1)$ ,  $\bar{\beta}_{n,0,\rho}(w) = 0$  and  $\bar{\beta}_{n,1,\rho}(w) = 1$ , it suffices to prove that, for any  $w \in (0, 1)$ ,  $\bar{\beta}_{n,t_1,\rho}(w) \leq \bar{\beta}_{n,t_2,\rho}(w)$  whenever  $0 < t_1 \leq t_2 < 1$ . Fix  $0 < t_1 \leq t_2 < 1$  and let  $X$  be Beta( $t_1\tau, (1-t_1)\tau$ ) and  $Y$  be Beta( $t_2\tau, (1-t_2)\tau$ ), where  $\tau = (n-\rho)/\rho > 0$ . As explained in the proof of Lemma 3.8.4, it then suffices to prove that  $X \leq_{lr} Y$ .

Let  $f$  be the density of  $X$  and  $g$  be the density of  $Y$ . By definition, we have that

$$\begin{aligned} f(x) &= \frac{\Gamma(\tau)}{\Gamma(t_1\tau)\Gamma\{(1-t_1)\tau\}} x^{t_1\tau-1} (1-x)^{(1-t_1)\tau-1}, \quad x \in (0, 1) \\ g(x) &= \frac{\Gamma(\tau)}{\Gamma(t_2\tau)\Gamma\{(1-t_2)\tau\}} x^{t_2\tau-1} (1-x)^{(1-t_2)\tau-1}, \quad x \in (0, 1). \end{aligned}$$

Accordinging for instance to Section 1.C.1 in Shaked and Shanthikumar (2007), to prove that  $X \leq_{lr} Y$ , we need to show that the function  $x \mapsto g(x)/f(x)$  is increasing on  $(0, 1)$ . Let  $l(x) = \log\{g(x)/f(x)\}$ ,  $x \in (0, 1)$ . Then, we have that, for any  $x \in (0, 1)$ ,

$$\begin{aligned} \frac{d}{dx}\{g(x)/f(x)\} &= \frac{d}{dx}\{e^{l(x)}\} = e^{l(x)} \frac{d}{dx}\{l(x)\} \\ &= e^{l(x)} \frac{d}{dx} \left[ \log \left\{ \frac{\frac{\Gamma(\tau)}{\Gamma(t_2\tau)\Gamma\{(1-t_2)\tau\}} x^{t_2\tau-1} (1-x)^{(1-t_2)\tau-1}}{\frac{\Gamma(\tau)}{\Gamma(t_1\tau)\Gamma\{(1-t_1)\tau\}} x^{t_1\tau-1} (1-x)^{(1-t_1)\tau-1}} \right\} \right] \\ &= e^{l(x)} \frac{d}{dx} \left[ \log \left\{ \frac{\Gamma(t_1\tau)\Gamma\{(1-t_1)\tau\}}{\Gamma(t_2\tau)\Gamma\{(1-t_2)\tau\}} \right\} + \log \left\{ \frac{x^{t_2\tau-1} (1-x)^{(1-t_2)\tau-1}}{x^{t_1\tau-1} (1-x)^{(1-t_1)\tau-1}} \right\} \right] \\ &= e^{l(x)} \frac{d}{dx} \{(t_2-t_1)\tau \log(x) - (t_2-t_1)\tau \log(1-x)\} \\ &= e^{l(x)} (t_2-t_1)\tau \left( \frac{1}{x} + \frac{1}{1-x} \right) \geq 0. \end{aligned}$$

■

**Lemma 3.8.6.** *For any  $n \in \mathbb{N}$ ,  $\rho \in (1, n)$  and  $w \in [0, n)$ , the function  $t \mapsto \bar{\mathcal{B}}_{n,t,\rho}(w)$  is increasing on  $[0, 1]$ .*

**Proof.** Fix  $n \in \mathbb{N}$  and  $\rho \in (1, n)$  and let  $\tau = (n - \rho)/(\rho - 1) > 0$ . We need to prove that, for any  $w \in [0, n)$ ,  $\bar{\mathcal{B}}_{n,t_1,\rho}(w) \leq \bar{\mathcal{B}}_{n,t_2,\rho}(w)$  whenever  $0 \leq t_1 \leq t_2 \leq 1$ . For any  $w \in [0, n)$ ,  $\bar{\mathcal{B}}_{n,0,\rho}(w) = 0$  and  $\bar{\mathcal{B}}_{n,1,\rho}(w) = 1$  so that it suffices to prove that, for any  $w \in [0, n)$ ,  $\bar{\mathcal{B}}_{n,t_1,\rho}(w) \leq \bar{\mathcal{B}}_{n,t_2,\rho}(w)$  whenever  $0 < t_1 \leq t_2 < 1$ . Fix  $0 < t_1 \leq t_2 < 1$  and let  $X$  be Beta-Binomial( $n, t_1\tau, (1 - t_1)\tau$ ) and  $Y$  be Beta-Binomial( $n, t_2\tau, (1 - t_2)\tau$ ). The latter is then proven if we show that  $X \leq_{\text{st}} Y$ .

It is well-known that the distributions of  $X$  and  $Y$  can be viewed as compound distributions: let  $\Theta_1$  (resp.  $\Theta_2$ ) be Beta( $t_1\tau, (1 - t_1)\tau$ ) (resp. Beta( $t_2\tau, (1 - t_2)\tau$ )) and let  $Z_p$  be Binomial( $n, p$ ); then, with some abuse of notation,  $X$  (resp.  $Y$ ) has the same distribution as  $Z_{\Theta_1}$  (resp.  $Z_{\Theta_2}$ ). From the proof of Lemma 3.8.4, we have that  $Z_{p_1} \leq_{\text{st}} Z_{p_2}$  whenever  $0 < p_1 \leq p_2 < 1$ , while from the proof of Lemma 3.8.5, we have that  $\Theta_1 \leq_{\text{st}} \Theta_2$ . Theorem 1.A.6 in Shaked and Shanthikumar (2007) then implies that  $X \leq_{\text{st}} Y$ , which completes the proof. ■

**Proofs of Propositions 3.4.1, 3.4.2 and 3.4.4.** Proposition 3.4.1 (resp. Proposition 3.4.2, Proposition 3.4.4) is an immediate consequence of Lemmas 3.8.1 and 3.8.4 (resp. Lemmas 3.8.3 and 3.8.6, Lemmas 3.8.2 and 3.8.5). ■

### 3.8.3 Proofs of Theorem 3.6.4

The proof of Theorem 3.6.4 is based on two lemmas which can be seen as extensions of similar results stated in Section 3 of Segers et al. (2017). The first lemma involves the following condition which is implied by Condition 3.6.2.

**Condition 3.8.7.** There exists a positive sequence  $h_n \downarrow 0$  such that, for any  $n \in \mathbb{N}$ ,  $\mathbf{x} \in (\mathbb{R}^d)^n$ ,  $\mathbf{u} \in [0, 1]^d$  and  $j \in \{1, \dots, d\}$ ,  $\text{Var}(W_{j,u_j}^{\mathbf{x}}) \leq h_n$ .

**Lemma 3.8.8.** *Let  $\mathbb{X}_n$  be a process in  $\ell^\infty(\Delta \times [0, 1]^d)$  such that, for all  $\mathbf{u} \in [0, 1]^d$  and  $s \in [0, 1]$ ,  $\mathbb{X}_n(s, s, \mathbf{u}) = 0$ . Furthermore, assume that  $\mathbb{X}_n \rightsquigarrow \mathbb{X}$  in  $\ell^\infty(\Delta \times [0, 1]^d)$  where  $\mathbb{X}$  has continuous trajectories almost surely. Then, under Condition 3.8.7,*

$$\sup_{(s,t,\mathbf{u}) \in \Delta \times [0,1]^d} \left| \int_{[0,1]^d} \mathbb{X}_n(s, t, \mathbf{w}) d\nu_{\mathbf{u}}^{\mathbf{x}^{\lfloor ns \rfloor + 1: \lfloor nt \rfloor}}(\mathbf{w}) - \mathbb{X}_n(s, t, \mathbf{u}) \right| = o_{\mathbb{P}}(1).$$

**Proof.** Let  $\|\cdot\|_\infty$  denote the maximum norm on  $\mathbb{R}^d$ . For any  $\varepsilon > 0$ ,

$$\begin{aligned}
 & \sup_{\substack{(s,t) \in \Delta \\ \mathbf{u} \in [0,1]^d}} \left| \int_{[0,1]^d} \mathbb{X}_n(s,t,\mathbf{w}) d\nu_{\mathbf{u}}^{\mathcal{X}^{[ns]+1:[nt]}}(\mathbf{w}) - \mathbb{X}_n(s,t,\mathbf{u}) \right| \\
 &= \sup_{\substack{(s,t) \in \Delta \\ \mathbf{u} \in [0,1]^d}} \left| \int_{[0,1]^d} \{\mathbb{X}_n(s,t,\mathbf{w}) - \mathbb{X}_n(s,t,\mathbf{u})\} d\nu_{\mathbf{u}}^{\mathcal{X}^{[ns]+1:[nt]}}(\mathbf{w}) \right| \\
 &\leq \sup_{\substack{(s,t) \in \Delta \\ \mathbf{u} \in [0,1]^d}} \left| \int_{\{\mathbf{w} \in [0,1]^d : |\mathbf{u}-\mathbf{w}|_\infty \leq \varepsilon\}} \{\mathbb{X}_n(s,t,\mathbf{w}) - \mathbb{X}_n(s,t,\mathbf{u})\} d\nu_{\mathbf{u}}^{\mathcal{X}^{[ns]+1:[nt]}}(\mathbf{w}) \right| \\
 &\quad + \sup_{\substack{(s,t) \in \Delta \\ \mathbf{u} \in [0,1]^d}} \left| \int_{\{\mathbf{w} \in [0,1]^d : |\mathbf{u}-\mathbf{w}|_\infty > \varepsilon\}} \{\mathbb{X}_n(s,t,\mathbf{w}) - \mathbb{X}_n(s,t,\mathbf{u})\} d\nu_{\mathbf{u}}^{\mathcal{X}^{[ns]+1:[nt]}}(\mathbf{w}) \right| \\
 &\leq \sup_{\substack{(s,t,\mathbf{u},\mathbf{w}) \in \Delta \times [0,1]^{2d} \\ |\mathbf{u}-\mathbf{w}|_\infty \leq \varepsilon}} |\mathbb{X}_n(s,t,\mathbf{w}) - \mathbb{X}_n(s,t,\mathbf{u})| + \sup_{\substack{(s,t) \in \Delta \\ (\mathbf{u},\mathbf{w}) \times [0,1]^{2d}}} \left[ |\mathbb{X}_n(s,t,\mathbf{w}) - \mathbb{X}_n(s,t,\mathbf{u})| \right. \\
 &\quad \left. \times \nu_{\mathbf{u}}^{\mathcal{X}^{[ns]+1:[nt]}}(\{\mathbf{w} \in [0,1]^d : |\mathbf{u}-\mathbf{w}|_\infty > \varepsilon\}) \right].
 \end{aligned}$$

Since  $\mathbb{X}_n \rightsquigarrow \mathbb{X}$  in  $\ell^\infty(\Delta \times [0,1]^d)$  and  $\mathbb{X}$  has continuous trajectories almost surely, it is stochastically equicontinuous. Hence, for any given  $\eta > 0$ , we can choose  $\varepsilon = \varepsilon(\eta) > 0$  sufficiently small such that

$$\limsup_{n \rightarrow \infty} \mathbb{P} \left\{ \sup_{\substack{(s,t,\mathbf{u},\mathbf{w}) \in \Delta \times [0,1]^{2d} \\ |\mathbf{u}-\mathbf{w}|_\infty \leq \varepsilon}} |\mathbb{X}_n(s,t,\mathbf{w}) - \mathbb{X}_n(s,t,\mathbf{u})| > \eta \right\} \leq \eta.$$

Some thought then reveals that, to complete the proof, it suffices to show that

$$\sup_{\substack{(s,t) \in \Delta \\ (\mathbf{u},\mathbf{w}) \times [0,1]^{2d}}} \left[ |\mathbb{X}_n(s,t,\mathbf{w}) - \mathbb{X}_n(s,t,\mathbf{u})| \nu_{\mathbf{u}}^{\mathcal{X}^{[ns]+1:[nt]}}(\{\mathbf{w} \in [0,1]^d : |\mathbf{u}-\mathbf{w}|_\infty > \varepsilon\}) \right] = o_{\mathbb{P}}(1). \tag{3.34}$$

For any  $\delta > 0$ , the supremum on the left-hand side of (3.34) is smaller than

$$\begin{aligned}
 & \sup_{\substack{(s,t,\mathbf{u},\mathbf{w}) \in \Delta \times [0,1]^{2d} \\ t-s \leq \delta}} \left[ |\mathbb{X}_n(s,t,\mathbf{w}) - \mathbb{X}_n(s,t,\mathbf{u})| \times \nu_{\mathbf{u}}^{\mathcal{X}^{[ns]+1:[nt]}}(\{\mathbf{w} \in [0,1]^d : |\mathbf{u}-\mathbf{w}|_\infty > \varepsilon\}) \right] \\
 &+ \sup_{\substack{(s,t,\mathbf{u},\mathbf{w}) \in \Delta \times [0,1]^{2d} \\ t-s > \delta}} \left[ |\mathbb{X}_n(s,t,\mathbf{w}) - \mathbb{X}_n(s,t,\mathbf{u})| \times \nu_{\mathbf{u}}^{\mathcal{X}^{[ns]+1:[nt]}}(\{\mathbf{w} \in [0,1]^d : |\mathbf{u}-\mathbf{w}|_\infty > \varepsilon\}) \right],
 \end{aligned}$$

which is in turn smaller than

$$2 \sup_{\substack{(s,t,\mathbf{u}) \in \Delta \times [0,1]^d \\ t-s \leq \delta}} |\mathbb{X}_n(s,t,\mathbf{u})| + 2 \sup_{(s,t,\mathbf{u}) \in \Delta \times [0,1]^d} |\mathbb{X}_n(s,t,\mathbf{u})| \\ \times \sup_{\substack{(s,t,\mathbf{u}) \in \Delta \times [0,1]^d \\ t-s > \delta}} \nu_{\mathbf{u}}^{\mathbf{x}^{\lfloor ns \rfloor + 1 : \lfloor nt \rfloor}}(\{\mathbf{w} \in [0,1]^d : |\mathbf{u} - \mathbf{w}|_{\infty} > \varepsilon\}).$$

Using again the fact that  $\mathbb{X}_n$  is stochastically equicontinuous and that it vanishes on the subset  $\{(s,s,\mathbf{u}) : s \in [0,1], \mathbf{u} \in [0,1]^d\}$  of  $\Delta \times [0,1]^d$ , for any given  $\eta > 0$ , we can choose  $\delta = \delta(\eta) > 0$  sufficiently small such that

$$\limsup_{n \rightarrow \infty} \mathbb{P} \left\{ \sup_{\substack{(s,t,\mathbf{u}) \in \Delta \times [0,1]^d \\ t-s \leq \delta}} |\mathbb{X}_n(s,t,\mathbf{u})| > \eta \right\} \\ = \limsup_{n \rightarrow \infty} \mathbb{P} \left\{ \sup_{\substack{(s,t,\mathbf{u}) \in \Delta \times [0,1]^d \\ t-s \leq \delta}} |\mathbb{X}_n(s,t,\mathbf{u}) - \mathbb{X}_n(s,s,\mathbf{u})| > \eta \right\} \leq \eta.$$

The convergence in (3.34) will then hold if we additionally show that

$$\sup_{(s,t,\mathbf{u}) \in \Delta \times [0,1]^d} |\mathbb{X}_n(s,t,\mathbf{u})| \times \sup_{\substack{(s,t,\mathbf{u}) \in \Delta \times [0,1]^d \\ t-s > \delta}} \nu_{\mathbf{u}}^{\mathbf{x}^{\lfloor ns \rfloor + 1 : \lfloor nt \rfloor}}(\{\mathbf{w} \in [0,1]^d : |\mathbf{u} - \mathbf{w}|_{\infty} > \varepsilon\}) = o_{\mathbb{P}}(1). \quad (3.35)$$

From the weak convergence of  $\mathbb{X}_n$ , we have that  $\sup_{(s,t,\mathbf{u}) \in \Delta \times [0,1]^d} |\mathbb{X}_n(s,t,\mathbf{u})| = O_{\mathbb{P}}(1)$  whereas the second factor in the product on the left-hand side of (3.35) can be shown to converge to zero almost surely. Indeed, proceeding along the lines of (Segers et al., 2017, Section 3) using Chebyshev's inequality and Condition 3.8.7, for almost any sequence  $\mathbf{X}_1, \mathbf{X}_2, \dots$ , conditionally

on  $\mathbf{X}_1, \mathbf{X}_2, \dots$ , we obtain that

$$\begin{aligned}
 & \sup_{\substack{(s,t,\mathbf{u}) \in \Delta \times [0,1]^d \\ t-s > \delta}} \nu_{\mathbf{u}}^{\mathbf{x}_{[ns]+1:[nt]}}(\{\mathbf{w} \in [0,1]^d : |\mathbf{u} - \mathbf{w}|_{\infty} > \varepsilon\}) \\
 &= \sup_{\substack{(s,t,\mathbf{u}) \in \Delta \times [0,1]^d \\ t-s > \delta}} \mathbb{P} \left\{ \left| \mathbf{W}_{\mathbf{u}}^{\mathbf{x}_{[ns]+1:[nt]}} - \mathbf{u} \right|_{\infty} > \varepsilon \mid \mathbf{x}_{[ns]+1:[nt]} \right\} \\
 &= \sup_{\substack{(s,t,\mathbf{u}) \in \Delta \times [0,1]^d \\ t-s > \delta}} \mathbb{P} \left[ \bigcup_{j=1}^d \left\{ \left| W_{j,u_j}^{\mathbf{x}_{[ns]+1:[nt]}} - u_j \right| > \varepsilon \right\} \mid \mathbf{x}_{[ns]+1:[nt]} \right] \\
 &\leq \sup_{\substack{(s,t,\mathbf{u}) \in \Delta \times [0,1]^d \\ t-s > \delta}} \sum_{j=1}^d \mathbb{P} \left\{ \left| W_{j,u_j}^{\mathbf{x}_{[ns]+1:[nt]}} - u_j \right| > \varepsilon \mid \mathbf{x}_{[ns]+1:[nt]} \right\} \\
 &\leq \sup_{\substack{(s,t,\mathbf{u}) \in \Delta \times [0,1]^d \\ t-s > \delta}} \sum_{j=1}^d \frac{\text{Var}(W_{j,u_j}^{\mathbf{x}_{[ns]+1:[nt]}} \mid \mathbf{x}_{[ns]+1:[nt]})}{\varepsilon^2} \\
 &\leq \frac{d}{\varepsilon^2} \sup_{\substack{(s,t) \in \Delta \\ t-s > \delta}} h_{[nt]-[ns]} \leq \frac{d}{\varepsilon^2} \sup_{\substack{(s,t) \in \Delta \\ t-s > \delta}} h_{[n(t-s)-1]} \leq \frac{d}{\varepsilon^2} h_{[n\delta-1]} = o(1).
 \end{aligned}$$

It follows that the second factor in the product on the left-hand side of (3.35) converges almost surely to zero, which completes the proof.  $\blacksquare$

**Lemma 3.8.9.** *Assume that Conditions 3.6.1 and 3.6.2 hold. Then, almost surely,*

$$\sup_{(s,t,\mathbf{u}) \in \Delta \times [0,1]^d} \sqrt{n} \lambda_n(s,t) \left| \int_{[0,1]^d} C(\mathbf{w}) d\nu_{\mathbf{u}}^{\mathbf{x}_{[ns]+1:[nt]}}(\mathbf{w}) - C(\mathbf{u}) \right| = o(1).$$

**Proof.** Let  $a_n = \lfloor n^{1/3} \rfloor$ ,  $n \in \mathbb{N}$ . With probability 1, one has that

$$\begin{aligned}
 & \sup_{(s,t,\mathbf{u}) \in \Delta \times [0,1]^d} \sqrt{n} \lambda_n(s,t) \left| \int_{[0,1]^d} C(\mathbf{w}) d\nu_{\mathbf{u}}^{\mathbf{x}_{[ns]+1:[nt]}}(\mathbf{w}) - C(\mathbf{u}) \right| \\
 &= \max_{1 \leq k \leq l \leq n} \left\{ \frac{l-k+1}{\sqrt{n}} \sup_{\mathbf{u} \in [0,1]^d} \left| \int_{[0,1]^d} C(\mathbf{w}) d\nu_{\mathbf{u}}^{\mathbf{x}_{k:l}}(\mathbf{w}) - C(\mathbf{u}) \right| \right\} \leq \max(L_n, M_n),
 \end{aligned}$$

where

$$\begin{aligned}
 L_n &= \max_{\substack{1 \leq k \leq l \leq n \\ l-k \leq a_n}} \left\{ \frac{l-k+1}{\sqrt{n}} \sup_{\mathbf{u} \in [0,1]^d} \left| \int_{[0,1]^d} C(\mathbf{w}) d\nu_{\mathbf{u}}^{\mathbf{x}_{k:l}}(\mathbf{w}) - C(\mathbf{u}) \right| \right\}, \\
 M_n &= \max_{\substack{1 \leq k \leq l \leq n \\ l-k > a_n}} \left\{ \sqrt{l-k+1} \sup_{\mathbf{u} \in [0,1]^d} \left| \int_{[0,1]^d} C(\mathbf{w}) d\nu_{\mathbf{u}}^{\mathbf{x}_{k:l}}(\mathbf{w}) - C(\mathbf{u}) \right| \right\}. \tag{3.36}
 \end{aligned}$$

Since  $0 \leq C \leq 1$ , we have that  $\sup_{\mathbf{u} \in [0,1]^d} \left| \int_{[0,1]^d} C(\mathbf{w}) d\nu_{\mathbf{u}}^{\mathbf{x}^{k:l}}(\mathbf{w}) - C(\mathbf{u}) \right| \leq 1$  almost surely for all  $1 \leq k \leq l \leq n$  and, therefore, with probability 1, that

$$L_n \leq \max_{\substack{1 \leq k \leq l \leq n \\ l-k \leq a_n}} \frac{l-k+1}{\sqrt{n}} \leq \frac{a_n+1}{\sqrt{n}} = o(1).$$

The aim of the remainder of this proof is to show that the term  $M_n$  in (3.36) converges to zero almost surely as well. To do so, it suffices to show that, for almost any sequence  $\mathbf{X}_1, \mathbf{X}_2, \dots$ , conditionally on  $\mathbf{X}_1, \mathbf{X}_2, \dots$ ,  $M_n$  converges to zero. We thus reason conditionally on  $\mathbf{X}_1, \mathbf{X}_2, \dots$  in the rest of this proof.

Let  $\eta > 0$  and let us show that  $M_n$  can be made smaller than  $d\eta$  provided  $n$  is large enough. The forthcoming arguments are very close to those used in the proof of Proposition 3.4 in (Segers et al., 2017). We provide all the steps nonetheless for the sake of completeness.

Given  $\mathbf{u}, \mathbf{w} \in [0,1]^d$ , set  $\mathbf{w}(t) = \mathbf{u} + t(\mathbf{w} - \mathbf{u})$ ,  $t \in [0,1]$ . The function  $G(t) = C\{\mathbf{w}(t)\}$ ,  $t \in [0,1]$ , is continuous on  $[0,1]$  and, by Condition 3.6.1, is continuously differentiable on  $(0,1)$  with derivative

$$G'(t) = \sum_{j=1}^d (w_j - u_j) \dot{C}_j\{\mathbf{w}(t)\}, \quad t \in (0,1).$$

By the fundamental theorem of calculus,  $G(1) - G(0) = \int_0^1 G'(t) dt$ , that is,

$$C(\mathbf{w}) - C(\mathbf{u}) = \sum_{j=1}^d (w_j - u_j) \int_0^1 \dot{C}_j\{\mathbf{w}(t)\} dt.$$

Note that, under Condition 3.6.1 and with the adopted conventions, some thought reveals that the previous equality holds no matter how  $\mathbf{u}$  and  $\mathbf{w}$  are chosen in  $[0,1]^d$ . Using Fubini's theorem, we then obtain that, for any  $\mathbf{u} \in [0,1]^d$  and  $1 \leq k \leq l \leq n$ ,

$$\int_{[0,1]^d} \{C(\mathbf{w}) - C(\mathbf{u})\} d\nu_{\mathbf{u}}^{\mathbf{x}^{k:l}}(\mathbf{w}) = \sum_{j=1}^d \int_0^1 \left\{ \int_{[0,1]^d} (w_j - u_j) \dot{C}_j\{\mathbf{w}(t)\} d\nu_{\mathbf{u}}^{\mathbf{x}^{k:l}}(\mathbf{w}) \right\} dt,$$

which implies that

$$\begin{aligned} M_n &= \max_{\substack{1 \leq k \leq l \leq n \\ l-k > a_n}} \left\{ \sqrt{l-k+1} \sup_{\mathbf{u} \in [0,1]^d} \left| \sum_{j=1}^d \int_0^1 \left\{ \int_{[0,1]^d} (w_j - u_j) \dot{C}_j\{\mathbf{w}(t)\} d\nu_{\mathbf{u}}^{\mathbf{x}^{k:l}}(\mathbf{w}) \right\} dt \right| \right\} \\ &\leq \max_{\substack{1 \leq k \leq l \leq n \\ l-k > a_n}} \left\{ \sqrt{l-k+1} \sum_{j=1}^d \sup_{\mathbf{u} \in [0,1]^d} \left| \int_0^1 \left\{ \int_{[0,1]^d} (w_j - u_j) \dot{C}_j\{\mathbf{w}(t)\} d\nu_{\mathbf{u}}^{\mathbf{x}^{k:l}}(\mathbf{w}) \right\} dt \right| \right\} \\ &\leq \sum_{j=1}^d I_{j,n}, \end{aligned}$$

where, for any  $j \in \{1, \dots, d\}$ ,

$$I_{j,n} = \max_{\substack{1 \leq k \leq l \leq n \\ l-k > a_n}} \left\{ \sqrt{l-k+1} \int_0^1 \sup_{\mathbf{u} \in [0,1]^d} \left| \int_{[0,1]^d} (w_j - u_j) \dot{C}_j\{\mathbf{w}(t)\} d\nu_{\mathbf{u}}^{\mathbf{x}_{k:l}}(\mathbf{w}) \right| dt \right\}.$$

Fix  $j \in \{1, \dots, d\}$ . We shall now show that, provided  $n$  is large enough,  $I_{j,n}$  is smaller than  $\eta$ . For any  $\delta \in (0, 1/2]$ , we have that  $I_{j,n} \leq J_{j,n,\delta} + K_{j,n,\delta}$ , where

$$\begin{aligned} J_{j,n,\delta} &= \max_{\substack{1 \leq k \leq l \leq n \\ l-k > a_n}} \left\{ \sqrt{l-k+1} \int_0^1 \sup_{\substack{\mathbf{u} \in [0,1]^d \\ u_j \in [0,\delta) \cup (1-\delta,1]}} \left| \int_{[0,1]^d} (w_j - u_j) \dot{C}_j\{\mathbf{w}(t)\} d\nu_{\mathbf{u}}^{\mathbf{x}_{k:l}}(\mathbf{w}) \right| dt \right\}, \\ K_{j,n,\delta} &= \max_{\substack{1 \leq k \leq l \leq n \\ l-k > a_n}} \left\{ \sqrt{l-k+1} \int_0^1 \sup_{\substack{\mathbf{u} \in [0,1]^d \\ u_j \in [\delta,1-\delta]}} \left| \int_{[0,1]^d} (w_j - u_j) \dot{C}_j\{\mathbf{w}(t)\} d\nu_{\mathbf{u}}^{\mathbf{x}_{k:l}}(\mathbf{w}) \right| dt \right\}. \end{aligned} \quad (3.37)$$

*Term  $J_{j,n,\delta}$ :* From the monotonicity and Lipschitz continuity of  $C$ ,  $0 \leq \dot{C}_j \leq 1$  (see, e.g., [Nelsen, 2006](#), Section 2.2) and therefore, using additionally Hölder's inequality and Condition [3.6.2](#),

$$\begin{aligned} J_{j,n,\delta} &\leq \max_{\substack{1 \leq k \leq l \leq n \\ l-k > a_n}} \left\{ \sqrt{l-k+1} \sup_{\substack{\mathbf{u} \in [0,1]^d \\ u_j \in [0,\delta) \cup (1-\delta,1]}} \int_{[0,1]^d} |w_j - u_j| d\nu_{\mathbf{u}}^{\mathbf{x}_{k:l}}(\mathbf{w}) \right\} \\ &\leq \max_{\substack{1 \leq k \leq l \leq n \\ l-k > a_n}} \left\{ \sqrt{l-k+1} \sup_{\substack{\mathbf{u} \in [0,1]^d \\ u_j \in [0,\delta) \cup (1-\delta,1]}} \sqrt{\int_{[0,1]^d} \{w_j - \mathbb{E}(W_{j,u_j}^{\mathbf{x}_{k:l}} | \mathbf{x}_{k:l})\}^2 d\nu_{\mathbf{u}}^{\mathbf{x}_{k:l}}(\mathbf{w})} \right\} \\ &= \max_{\substack{1 \leq k \leq l \leq n \\ l-k > a_n}} \left\{ \sqrt{l-k+1} \sup_{\substack{\mathbf{u} \in [0,1]^d \\ u_j \in [0,\delta) \cup (1-\delta,1]}} \sqrt{\text{Var}(W_{j,u_j}^{\mathbf{x}_{k:l}} | \mathbf{x}_{k:l})} \right\} \leq \sup_{u \in [0,\delta) \cup (1-\delta,1]} \sqrt{\kappa u(1-u)}. \end{aligned}$$

Hence, we can choose  $\delta = \delta(\eta, \kappa)$  sufficiently small such that  $J_{j,n,\delta} < \eta/3$ .

*Term  $K_{j,n,\delta}$ :* Since  $\int_{[0,1]^d} (w_j - u_j) d\nu_{\mathbf{u}}^{\mathbf{x}_{k:l}}(\mathbf{w}) = 0$  for all  $1 \leq k \leq l \leq n$ , we can rewrite the term  $K_{j,n,\delta}$  in [\(3.37\)](#) as

$$K_{j,n,\delta} = \max_{\substack{1 \leq k \leq l \leq n \\ l-k > a_n}} \left\{ \sqrt{l-k+1} \int_0^1 \sup_{\substack{\mathbf{u} \in [0,1]^d \\ u_j \in [\delta,1-\delta]}} \left| \int_{[0,1]^d} (w_j - u_j) \left[ \dot{C}_j\{\mathbf{w}(t)\} - \dot{C}_j(\mathbf{u}) \right] d\nu_{\mathbf{u}}^{\mathbf{x}_{k:l}}(\mathbf{w}) \right| dt \right\}.$$

Then, for any  $\varepsilon \in (0, \delta/2)$ ,



$$\begin{aligned}
 K_{j,n,\delta} &\leq \max_{\substack{1 \leq k \leq l \leq n \\ l-k > a_n}} \left\{ \sqrt{l-k+1} \int_0^1 \left[ \sup_{\substack{\mathbf{u} \in [0,1]^d \\ u_j \in [\delta, 1-\delta]}} \left| \int_{\{\mathbf{w} \in [0,1]^d : |\mathbf{w}-\mathbf{u}|_\infty \leq \varepsilon\}} (w_j - u_j) [\dot{C}_j\{\mathbf{w}(t)\} - \dot{C}_j(\mathbf{u})] d\nu_{\mathbf{u}}^{\mathbf{x}_{k:l}}(\mathbf{w}) \right| \right] dt \right\} \\
 &\quad + \max_{\substack{1 \leq k \leq l \leq n \\ l-k > a_n}} \left\{ \sqrt{l-k+1} \int_0^1 \left[ \sup_{\substack{\mathbf{u} \in [0,1]^d \\ u_j \in [\delta, 1-\delta]}} \left| \int_{\{\mathbf{w} \in [0,1]^d : |\mathbf{w}-\mathbf{u}|_\infty > \varepsilon\}} (w_j - u_j) [\dot{C}_j\{\mathbf{w}(t)\} - \dot{C}_j(\mathbf{u})] d\nu_{\mathbf{u}}^{\mathbf{x}_{k:l}}(\mathbf{w}) \right| \right] dt \right\} \\
 &\leq K'_{j,n,\delta,\varepsilon} + K''_{j,n,\delta,\varepsilon},
 \end{aligned}$$

where

$$\begin{aligned}
 K'_{j,n,\delta,\varepsilon} &= \sup_{\substack{\mathbf{u} \in [0,1]^d \\ u_j \in [\delta, 1-\delta]}} \sup_{\substack{\mathbf{w} \in [0,1]^d \\ |\mathbf{w}-\mathbf{u}|_\infty \leq \varepsilon}} |\dot{C}_j(\mathbf{w}) - \dot{C}_j(\mathbf{u})| \times \max_{\substack{1 \leq k \leq l \leq n \\ l-k > a_n}} \left\{ \sqrt{l-k+1} \sup_{\substack{\mathbf{u} \in [0,1]^d \\ u_j \in [\delta, 1-\delta]}} \int_{[0,1]^d} |w_j - u_j| d\nu_{\mathbf{u}}^{\mathbf{x}_{k:l}}(\mathbf{w}) \right\}, \\
 K''_{j,n,\delta,\varepsilon} &= \max_{\substack{1 \leq k \leq l \leq n \\ l-k > a_n}} \left\{ \sqrt{l-k+1} \sup_{\substack{\mathbf{u} \in [0,1]^d \\ u_j \in [\delta, 1-\delta]}} \int_{[0,1]^d} \mathbf{1}\{|\mathbf{w}-\mathbf{u}|_\infty > \varepsilon\} |w_j - u_j| d\nu_{\mathbf{u}}^{\mathbf{x}_{k:l}}(\mathbf{w}) \right\},
 \end{aligned}$$

where we have used the fact that, for all  $\mathbf{u} \in [0, 1]^d$  and  $t \in (0, 1)$ ,

$$\sup_{\substack{\mathbf{w} \in [0,1]^d \\ |\mathbf{w}-\mathbf{u}|_\infty \leq \varepsilon}} |\dot{C}_j\{\mathbf{w}(t)\} - \dot{C}_j(\mathbf{u})| \leq \sup_{\substack{\mathbf{w} \in [0,1]^d \\ |\mathbf{w}-\mathbf{u}|_\infty \leq \varepsilon}} |\dot{C}_j(\mathbf{w}) - \dot{C}_j(\mathbf{u})|$$

and, again, that  $0 \leq \dot{C}_j \leq 1$ . Using Condition 3.6.2, we obtain that, for any  $1 \leq k \leq l \leq n$ ,

$$\sup_{\substack{\mathbf{u} \in [0,1]^d \\ u_j \in [\delta, 1-\delta]}} \int_{[0,1]^d} |w_j - u_j| d\nu_{\mathbf{u}}^{\mathbf{x}_{k:l}}(\mathbf{w}) \leq \sup_{\substack{\mathbf{u} \in [0,1]^d \\ u_j \in [\delta, 1-\delta]}} \sqrt{\text{Var}(W_{j,u_j}^{\mathbf{x}_{k:l}} | \mathbf{x}_{k:l})} \leq \sup_{u \in [0,1]} \sqrt{\frac{\kappa u(1-u)}{l-k+1}} \leq \sqrt{\frac{\kappa}{l-k+1}}. \quad (3.38)$$

Since Condition 3.6.1 holds, by uniform continuity of  $\dot{C}_j$  on the set  $\{\mathbf{u} \in [0, 1]^d : u_j \in [\delta/2, 1 - \delta/2]\}$ , we can choose  $\varepsilon = \varepsilon(\delta, \eta, \kappa) > 0$  sufficiently small such that

$$\sup_{\substack{\mathbf{u} \in [0,1]^d \\ u_j \in [\delta, 1-\delta]}} \sup_{\substack{\mathbf{w} \in [0,1]^d \\ |\mathbf{w}-\mathbf{u}|_\infty \leq \varepsilon}} |\dot{C}_j(\mathbf{w}) - \dot{C}_j(\mathbf{u})| \leq \frac{\eta}{3\sqrt{\kappa}}. \quad (3.39)$$

Combining (3.38) and (3.39), we obtain that  $K'_{j,n,\delta,\varepsilon} \leq \eta/3$ . As far as  $K''_{j,n,\delta,\varepsilon}$  is concerned, using the Cauchy-Schwarz inequality,

$$\begin{aligned} K''_{j,n,\delta,\varepsilon} &\leq \max_{\substack{1 \leq k \leq l \leq n \\ l-k > a_n}} \left\{ \sqrt{l-k+1} \sup_{\substack{\mathbf{u} \in [0,1]^d \\ u_j \in [\delta,1-\delta]}} \sqrt{\int_{[0,1]^d} \mathbf{1}\{|\mathbf{w}-\mathbf{u}|_\infty > \varepsilon\} d\nu_{\mathbf{u}}^{\mathbf{x}_{k:l}}(\mathbf{w}) \int_{[0,1]^d} (w_j - u_j)^2 d\nu_{\mathbf{u}}^{\mathbf{x}_{k:l}}(\mathbf{w})} \right\} \\ &\leq \max_{\substack{1 \leq k \leq l \leq n \\ l-k > a_n}} \left\{ \sqrt{l-k+1} \sup_{\substack{\mathbf{u} \in [0,1]^d \\ u_j \in [\delta,1-\delta]}} \sqrt{\nu_{\mathbf{u}}^{\mathbf{x}_{k:l}}(\{\mathbf{w} \in [0,1]^d : |\mathbf{w}-\mathbf{u}|_\infty > \varepsilon\})} \sup_{\substack{\mathbf{u} \in [0,1]^d \\ u_j \in [\delta,1-\delta]}} \sqrt{\text{Var}(W_{j,u_j}^{\mathbf{x}_{k:l}} | \mathbf{x}_{k:l})} \right\}. \end{aligned} \quad (3.40)$$

Then, using Chebyshev's inequality and Condition 3.6.2, we obtain that, for any  $\mathbf{u} \in [0,1]^d$  such that  $u_j \in [\delta,1-\delta]$  and for any  $1 \leq k \leq l \leq n$ ,

$$\begin{aligned} \nu_{\mathbf{u}}^{\mathbf{x}_{k:l}}(\{\mathbf{w} \in [0,1]^d : |\mathbf{u}-\mathbf{w}|_\infty > \varepsilon\}) &= \mathbb{P} \left[ \bigcup_{j=1}^d \{ |W_{j,u_j}^{\mathbf{x}_{k:l}} - u_j| > \varepsilon \} \mid \mathbf{x}_{k:l} \right] \leq \sum_{j=1}^d \mathbb{P} \{ |W_{j,u_j}^{\mathbf{x}_{k:l}} - u_j| > \varepsilon \mid \mathbf{x}_{k:l} \} \\ &\leq \sum_{j=1}^d \frac{\text{Var}(W_{j,u_j}^{\mathbf{x}_{k:l}} | \mathbf{x}_{k:l})}{\varepsilon^2} \leq \frac{\kappa d}{(l-k+1)\varepsilon^2}. \end{aligned} \quad (3.41)$$

Furthermore, using Condition 3.6.2, for any  $1 \leq k \leq l \leq n$ ,

$$\sup_{\substack{\mathbf{u} \in [0,1]^d \\ u_j \in [\delta,1-\delta]}} \sqrt{\text{Var}(W_{j,u_j}^{\mathbf{x}_{k:l}} | \mathbf{x}_{k:l})} \leq \sup_{\substack{\mathbf{u} \in [0,1]^d \\ u_j \in [\delta,1-\delta]}} \sqrt{\frac{\kappa u_j (1-u_j)}{l-k+1}} \leq \sqrt{\frac{\kappa}{l-k+1}} \quad (3.42)$$

Hence, from (3.40), (3.41) and (3.42), we obtain that

$$\begin{aligned} K''_{j,n,\delta,\varepsilon} &\leq \max_{\substack{1 \leq k \leq l \leq n \\ l-k > a_n}} \left\{ \sqrt{l-k+1} \times \sqrt{\frac{\kappa d}{(l-k+1)\varepsilon^2}} \times \sqrt{\frac{\kappa}{l-k+1}} \right\} \\ &\leq \frac{\kappa\sqrt{d}}{\varepsilon} \times \max_{\substack{1 \leq k \leq l \leq n \\ l-k > a_n}} \sqrt{\frac{1}{l-k+1}} \leq \frac{\kappa\sqrt{d}}{\varepsilon} \times \sqrt{\frac{1}{a_n+1}}, \end{aligned}$$

which implies that, for  $n$  sufficiently large,  $K''_{j,n,\delta,\varepsilon} < \eta/3$ . Thus, provided that  $n$  is large enough, this successively implies that  $K_{j,n,\delta} \leq 2\eta/3$ , that  $I_{j,n} \leq \eta$  and finally, since  $M_n \leq \sum_{j=1}^d I_{j,n}$ , that  $M_n \leq d\eta$ . The latter holds conditionally on  $\mathbf{X}_1, \mathbf{X}_2, \dots$  for almost any sequence  $\mathbf{X}_1, \mathbf{X}_2, \dots$ , which completes the proof.  $\blacksquare$

**Proof of Theorem 3.6.4.** From (3.29) and (3.31), for any  $(s, t) \in \Delta$  and  $\mathbf{u} \in [0,1]^d$ , we have that

$$\begin{aligned} \mathbb{C}_n^\nu(s, t, \mathbf{u}) &= \sqrt{n} \lambda_n(s, t) \left[ \int_{[0,1]^d} \{C_{\lfloor ns \rfloor + 1 : \lfloor nt \rfloor}(\mathbf{w}) - C(\mathbf{w}) + C(\mathbf{w})\} d\nu_{\mathbf{u}}^{\mathbf{x}_{\lfloor ns \rfloor + 1 : \lfloor nt \rfloor}}(\mathbf{w}) - C(\mathbf{u}) \right] \\ &= \int_{[0,1]^d} \mathbb{C}_n(s, t, \mathbf{w}) d\nu_{\mathbf{u}}^{\mathbf{x}_{\lfloor ns \rfloor + 1 : \lfloor nt \rfloor}}(\mathbf{w}) + \sqrt{n} \lambda_n(s, t) \left\{ \int_{[0,1]^d} C(\mathbf{w}) d\nu_{\mathbf{u}}^{\mathbf{x}_{\lfloor ns \rfloor + 1 : \lfloor nt \rfloor}}(\mathbf{w}) - C(\mathbf{u}) \right\}. \end{aligned} \quad (3.43)$$

Following Segers et al. (2017), the first (resp. second) term on the right of (3.43) can be called the stochastic (resp. bias) term in the decomposition of  $\mathbb{C}_n^\nu(s, t, \mathbf{u})$ . Notice that both terms are equal to zero for any  $(s, t) \in \Delta$  such that  $\lfloor ns \rfloor = \lfloor nt \rfloor$ . It then immediately follows from the triangular inequality that

$$\begin{aligned} \sup_{\substack{(s,t) \in \Delta \\ \mathbf{u} \in [0,1]^d}} |\mathbb{C}_n^\nu(s, t, \mathbf{u}) - \mathbb{C}_n(s, t, \mathbf{u})| &\leq \sup_{\substack{(s,t) \in \Delta \\ \mathbf{u} \in [0,1]^d}} \left| \int_{[0,1]^d} \mathbb{C}_n(s, t, \mathbf{w}) d\nu_{\mathbf{u}}^{\mathbf{x}^{\lfloor ns \rfloor + 1; \lfloor nt \rfloor}}(\mathbf{w}) - \mathbb{C}_n(s, t, \mathbf{u}) \right| \\ &\quad + \sup_{\substack{(s,t) \in \Delta \\ \mathbf{u} \in [0,1]^d}} \sqrt{n} \lambda_n(s, t) \left| \int_{[0,1]^d} C(\mathbf{w}) d\nu_{\mathbf{u}}^{\mathbf{x}^{\lfloor ns \rfloor + 1; \lfloor nt \rfloor}}(\mathbf{w}) - C(\mathbf{u}) \right|, \end{aligned}$$

and the desired result is finally a direct consequence of Lemmas 3.8.8 and 3.8.9. ■

## Acknowledgments

The authors would like to thank an Associate Editor and two anonymous Referees for their very constructive comments on an earlier version of this manuscript as well as Mark Holmes for fruitful discussions.

# Chapter 4

## Resampling techniques for a class of smooth, possibly data-adaptive empirical copulas

This chapter is based on [Kojadinovic and Yi \(2023\)](https://arxiv.org/abs/2301.05495): Resampling techniques for a class of smooth, possibly data-adaptive empirical copulas. preprint. <https://arxiv.org/abs/2301.05495>.

4.1	Introduction . . . . .	76
4.2	Smooth, possibly data-adaptive, empirical copulas and their asymptotics . . . . .	79
4.2.1	Smooth, possibly data-adaptive, nonparametric copula estimators . . . . .	79
4.2.2	Asymptotics of related sequential processes . . . . .	82
4.3	Bootstrap by drawing samples from the estimators in the i.i.d. case . . . . .	84
4.3.1	Drawing samples from the smooth empirical copulas . . . . .	85
4.3.2	Asymptotic validity results . . . . .	87
4.3.3	Finite-sample comparison of two smooth bootstraps . . . . .	88
4.4	Smooth sequential dependent multiplier bootstraps in the time series case . . . . .	90
4.4.1	Main intuition and existing work . . . . .	90
4.4.2	I.i.d. and dependent multiplier sequences . . . . .	91
4.4.3	Non-smooth sequential dependent multiplier replicates . . . . .	91
4.4.4	Smooth sequential dependent multiplier replicates . . . . .	92
4.4.5	Finite-sample comparison of three multiplier bootstraps . . . . .	94
4.4.6	Application to change-point detection . . . . .	99

4.5	Estimators of the first-order partial derivatives of the copula . . . . .	102
4.5.1	Estimators based on finite differences of the empirical copula . . . . .	103
4.5.2	Two classes of smooth estimators . . . . .	104
4.5.3	Weak consistency . . . . .	106
4.5.4	Finite-sample performance of selected estimators . . . . .	108
4.6	Conclusion . . . . .	112
4.7	Proofs . . . . .	113
4.7.1	Proof of Corollary 4.2.13 . . . . .	113
4.7.2	Proofs of Proposition 4.3.3 and Lemma 4.7.4 . . . . .	114
4.7.3	Proof of Theorem 4.3.5 . . . . .	116
4.7.4	Proof of Theorem 4.4.3 . . . . .	120
4.7.5	Proofs of Propositions 4.5.3, 4.5.5, 4.5.9 and 4.5.10 . . . . .	122

## Abstract

We investigate the validity of two resampling techniques when carrying out inference on the underlying unknown copula using a recently proposed class of smooth, possibly data-adaptive nonparametric estimators that contains empirical Bernstein copulas (and thus the empirical beta copula). Following [Kiriliouk et al. \(2021\)](#), the first resampling technique is based on drawing samples from the smooth estimator and can only be used in the case of independent observations. The second technique is a smooth extension of the so-called sequential dependent multiplier bootstrap and can thus be used in a time series setting and, possibly, for change-point analysis. The two studied resampling schemes are applied to confidence interval construction and the offline detection of changes in the cross-sectional dependence of multivariate time series, respectively. Monte Carlo experiments confirm the possible advantages of such smooth inference procedures over their non-smooth counterparts. A by-product of this work is the study of the weak consistency and finite-sample performance of two classes of smooth estimators of the first-order partial derivatives of a copula which can have applications in mean and quantile regression.

## 4.1 Introduction

Let  $\boldsymbol{x}_{1:n} = (\boldsymbol{X}_1, \dots, \boldsymbol{X}_n)$  be a stretch from a  $d$ -dimensional stationary time series  $(\boldsymbol{X}_i)_{i \in \mathbb{Z}}$  of continuous random vectors. From a well-known theorem due to [Sklar \(1959\)](#), the multivariate

distribution function (d.f.)  $F$  of each  $\mathbf{X}_i$  can be expressed as

$$F(\mathbf{x}) = C\{F_1(x_1), \dots, F_d(x_d)\}, \quad \mathbf{x} \in \mathbb{R}^d, \quad (4.1)$$

in terms of a unique *copula*  $C$  and the univariate margins  $F_1, \dots, F_d$  of  $F$ . Representation (4.1) is at root of many applications in probability, statistics and related fields (see, e.g., Hofert et al., 2018, and the references therein) because it suggests that  $F$  can be modeled in two separate steps: the first (resp. second) step consists of estimating the univariate margins  $F_1, \dots, F_d$  (resp. the copula  $C$ ). This work is only concerned with the estimation of the copula.

Statistical inference on the unknown copula  $C$  frequently involves the use of a nonparametric estimator of  $C$ . The best-known one is the *empirical copula* (Rüschendorf, 1976; Deheuvels, 1979) which we shall define as the empirical d.f. of the multivariate ranks obtained from  $\mathcal{X}_{1:n}$  scaled by  $1/n$ . Note that the latter function is piecewise constant and cannot therefore be a genuine copula. A promising smooth nonparametric estimator of  $C$  that is a genuine copula when there are no ties in the components samples of  $\mathcal{X}_{1:n}$  and that displays substantially better small-sample performance than the empirical copula is the *empirical beta copula*. This estimator was proposed by Segers et al. (2017) and is a particular case of the *empirical Bernstein copula* studied by Sancetta and Satchell (2004) and Janssen et al. (2012) when all the underlying Bernstein polynomials have degree  $n$ . Building upon the work of Segers et al. (2017), Kojadinovic and Yi (2022) recently studied data-adaptive generalizations of the empirical beta copula that can perform even better in small samples.

Whatever nonparametric estimator of the unknown copula  $C$  in (4.1) is used in inference procedures, it is almost always necessary to rely on resampling techniques to compute corresponding confidence intervals or p-values. To approximate the “sampling distribution” of the classical empirical copula, a frequently used approach in the literature is the so-called *multiplier bootstrap* (see, e.g., Scaillet, 2005; Rémillard and Scaillet, 2009). When the random vectors in  $\mathcal{X}_{1:n}$  are independent and identically distributed (i.i.d.), Bücher and Dette (2010) found the latter resampling scheme to have better finite-sample properties than approaches consisting of adapting the empirical (multinomial) bootstrap. The multiplier bootstrap was extended to the time series and sequential settings in Bücher and Kojadinovic (2016) and Bücher et al. (2014).

One of the advantages of the empirical beta copula is that it is particularly easy to draw samples from it. The resulting *smooth bootstrap* that can be used to approximate the “sampling distribution” of the empirical beta copula was recently studied both theoretically and empirically in Kiriliouk et al. (2021). The Monte Carlo experiments reported therein reveal that it is a competitive alternative to the multiplier bootstrap while being substantially simpler to implement. One practical inconvenience however is that the aforementioned smooth bootstrap cannot be directly extended to the time series setting.

The first aim of this work is to obtain, in the i.i.d. case, a smooth bootstrap *à la* Kiriliouk et al. (2021) for the smooth, possibly data-adaptive, nonparametric estimators of the copula investigated in Kojadinovic and Yi (2022). The second aim is to propose smooth versions of the

dependent multiplier bootstrap that can be used to approximate the “sampling distribution” of the aforementioned estimators in a time series setting. Intuitively, one could expect that the resulting smooth inference procedures will perform better than corresponding non-smooth procedures in particular when the amount of data is low. Indeed, as already mentioned, it is when  $n$  is small that smooth copula estimators can substantially outperform rough estimators such as the classical empirical copula; see for instance the finite-sample experiments reported in Segers et al. (2017), Kiriliouk et al. (2021) or Kojadinovic and Yi (2022). Another situation where one could expect that the use of smooth estimators can be advantageous is when carrying out change-point detection. Indeed, statistics for change-point detection often involve the comparison of estimators computed from small subsets of observations. It is to be able to cover this application area that many of the theoretical investigations carried out in this work are of a sequential nature.

A by-product of this work is the study of the weak consistency and finite-sample performance of two classes of smooth estimators of the first-order partial derivatives of the unknown copula  $C$  in (4.1) as these are needed to carry out the dependent multiplier bootstrap. As explained for instance in Janssen et al. (2016), such estimators have applications in mean and quantile regression as they lead to estimators of the conditional distribution function. From a practical perspective, our investigations lead to the proposal of a smooth data-adaptive estimator of the first-order partial derivatives of  $C$  that substantially outperforms, among others, the Bernstein estimator considered in Janssen et al. (2016).

This paper is organized as follows. In the second section, we recall the definition of the broad class of smooth, possibly data adaptive, empirical copulas studied in Kojadinovic and Yi (2022) and the asymptotics of related sequential empirical processes. The third section is concerned with an extension of the smooth bootstrap of Kiriliouk et al. (2021) that can be used to approximate the “sampling distribution” of the aforementioned smooth estimators in the i.i.d. case. After investigating its asymptotic validity, results of finite-sample experiments comparing smooth bootstraps based on the empirical beta copula and on its data-adaptive extension suggested in Kojadinovic and Yi (2022) are reported. In Section 4.4, to be able to cover the time series setting, we propose natural smooth extensions of the sequential dependent multiplier bootstrap. After providing asymptotic validity results, we compare the finite-sample performance of various versions of the multiplier bootstrap and consider an application to the offline detection of changes in the cross-sectional dependence of multivariate time series. The latter confirms the possible advantages of smooth inference procedures over their non-smooth counterparts. The fifth section is devoted to the study of two classes of smooth estimators of the first-order partial derivatives of  $C$ : their weak consistency is investigated and the finite-sample performance of selected estimators is studied.

Unless stated otherwise, all convergences in the paper are as  $n \rightarrow \infty$ . Also, in the sequel, the arrow ‘ $\rightsquigarrow$ ’ denotes weak convergence in the sense of Definition 1.3.3 in van der Vaart and Wellner (2000) and, given a set  $T$ ,  $\ell^\infty(T)$  (resp.  $\mathcal{C}(T)$ ) represents the space of all bounded (resp. continuous) real-valued functions on  $T$  equipped with the uniform metric.

All the numerical experiments presented in the work were carried out using the R statistical environment (R Core Team, 2022) as well as its packages `copula` (Hofert et al., 2022) and `extraDistr` (Wolodzko, 2020).

## 4.2 Smooth, possibly data-adaptive, empirical copulas and their asymptotics

In this section, we start by defining the broad class of smooth, possibly data adaptive, empirical copulas studied in Kojadinovic and Yi (2022). We then recall the asymptotics of related sequential empirical processes established in the same reference.

### 4.2.1 Smooth, possibly data-adaptive, nonparametric copula estimators

Because the results to be stated in the next section are of a sequential nature, all the quantities hereafter are defined for a substretch  $\mathcal{X}_{k:l} = (\mathbf{X}_k, \dots, \mathbf{X}_l)$ ,  $1 \leq k \leq l \leq n$ , of the available data  $\mathcal{X}_{1:n} = (\mathbf{X}_1, \dots, \mathbf{X}_n)$ .

For any  $j \in \{1, \dots, d\}$ , let  $F_{k:l,j}$  be the empirical d.f. computed from the  $j$ th component subsample  $X_{kj}, \dots, X_{lj}$  of  $\mathcal{X}_{k:l}$ . Then,  $R_{ij}^{k:l} = (l - k + 1)F_{k:l,j}(X_{ij}) = \sum_{t=k}^l \mathbf{1}(X_{tj} \leq X_{ij})$  is the (maximal) rank of  $X_{ij}$  among  $X_{kj}, \dots, X_{lj}$ . Furthermore, let

$$\mathbf{R}_i^{k:l} = \left( R_{i1}^{k:l}, \dots, R_{id}^{k:l} \right) \quad \text{and} \quad \hat{U}_i^{k:l} = \frac{\mathbf{R}_i^{k:l}}{l - k + 1}, \quad i \in \{k, \dots, l\},$$

be the multivariate ranks (resp. multivariate scaled ranks) obtained from  $\mathcal{X}_{k:l}$ . Following Rüschendorf (1976), the empirical copula  $C_{k:l}$  of  $\mathcal{X}_{k:l}$  is then defined, for any  $\mathbf{u} = (u_1, \dots, u_d) \in [0, 1]^d$ , by

$$C_{k:l}(\mathbf{u}) = \frac{1}{l - k + 1} \sum_{i=k}^l \prod_{j=1}^d \mathbf{1} \left( \frac{R_{ij}^{k:l}}{l - k + 1} \leq u_j \right) = \frac{1}{l - k + 1} \sum_{i=k}^l \mathbf{1}(\hat{U}_i^{k:l} \leq \mathbf{u}), \quad (4.2)$$

where inequalities between vectors are to be understood componentwise.

As we continue, following Kojadinovic and Yi (2022), for any  $m \in \mathbb{N}$ ,  $\mathbf{x} \in (\mathbb{R}^d)^m$  and  $\mathbf{u} \in [0, 1]^d$ ,  $\nu_{\mathbf{u}}^{\mathbf{x}}$  is the law of a  $[0, 1]^d$ -valued mean  $\mathbf{u}$  random vector  $\mathbf{W}_{\mathbf{u}}^{\mathbf{x}}$ . Its components are denoted by  $W_{1,u_1}^{\mathbf{x}}, \dots, W_{d,u_d}^{\mathbf{x}}$  to indicate that the  $j$ th component depends on  $u_j$  but not on  $u_1, \dots, u_{j-1}, u_{j+1}, \dots, u_d$ . Let  $p \geq d$  be a fixed integer and let  $\mathbf{U}$  be a  $p$ -dimensional random vector whose components are independent and standard uniform. The following assumption was considered in Kojadinovic and Yi (2022) and is likely to be non-restrictive as discussed in Remark 3 therein.

**Condition 4.2.1** (Construction of smoothing random vectors). For any  $m \in \mathbb{N}$ ,  $\mathbf{x} \in (\mathbb{R}^d)^m$  and  $\mathbf{u} \in [0, 1]^d$ , there exists a function  $\mathcal{W}_{\mathbf{u}}^{\mathbf{x}} : [0, 1]^p \rightarrow [0, 1]^d$  such that  $\mathbf{W}_{\mathbf{u}}^{\mathbf{x}} = \mathcal{W}_{\mathbf{u}}^{\mathbf{x}}(\mathbf{U})$ .



To be able to define, for any  $n \in \mathbb{N}$ ,  $\mathcal{X}_{1:n}$  and, for any  $m \leq n$ , the random vectors  $\mathbf{W}_{\mathbf{u}}^{\mathbf{x}}$ ,  $\mathbf{x} \in (\mathbb{R}^d)^m$ ,  $\mathbf{u} \in [0, 1]^d$ , on the same probability space  $(\Omega, \mathcal{A}, \mathbb{P})$ , we assume a product structure, that is,  $\Omega = \Omega_0 \times \Omega_1 \times \dots$  with probability measure  $\mathbb{P} = \mathbb{P}_0 \otimes \mathbb{P}_1 \otimes \dots$ , where  $\mathbb{P}_i$  denotes the probability measure on  $\Omega_i$ , such that, for any  $\omega \in \Omega$ ,  $\mathcal{X}_{1:n}(\omega)$  only depends on the first coordinate of  $\omega$ ,  $\mathbf{U}(\omega)$  only depends on the second coordinate of  $\omega$  and potential “bootstrap weights” (to be introduced in Sections 4.3 and 4.4) only depend on one of the remaining coordinates of  $\omega$ , implying in particular that  $\mathcal{X}_{1:n}$ ,  $\mathbf{U}$  and potential bootstrap weights are independent. A broad class of smooth versions of  $C_{k:l}$  in (4.2), with possibly data-adaptive smoothing, is then given by

$$C_{k:l}^{\nu}(\mathbf{u}) = \int_{[0,1]^d} C_{k:l}(\mathbf{w}) d\nu_{\mathbf{u}}^{\mathcal{X}_{k:l}}(\mathbf{w}), \quad \mathbf{u} \in [0, 1]^d. \quad (4.3)$$

Intuitively, for a given  $\mathbf{u} \in [0, 1]^d$ ,  $C_{k:l}^{\nu}(\mathbf{u})$  can be thought of as a “weighted average” of  $C_{k:l}(\mathbf{w})$  for  $\mathbf{w}$  “in a neighborhood of  $\mathbf{u}$ ” according to the smoothing distribution  $\nu_{\mathbf{u}}^{\mathcal{X}_{k:l}}$  (that may depend on the observations  $\mathcal{X}_{k:l}$ ). Note that, if  $k > l$ , we adopt the convention that  $C_{k:l} = C_{k:l}^{\nu} = 0$  and that, for any  $\mathbf{u} \in [0, 1]^d$ ,  $\nu_{\mathbf{u}}^{\mathcal{X}_{k:l}}$  is the Dirac measure at  $\mathbf{u}$ .

*Remark 4.2.2.* Given  $m \in \mathbb{N}$  and  $\mathbf{u} \in [0, 1]^d$ , let  $\mu_{m,\mathbf{u}}$  be the law of the  $d$ -dimensional random vector  $(S_{m,1,u_1}/m, \dots, S_{m,d,u_d}/m)$  such that the random variables  $S_{m,1,u_1}, \dots, S_{m,d,u_d}$  are independent and, for each  $j \in \{1, \dots, d\}$ ,  $S_{m,j,u_j}$  is Binomial( $m, u_j$ ). From Section 3 of Segers et al. (2017), the empirical Bernstein copula of  $\mathcal{X}_{k:l}$  whose Bernstein polynomial degrees are all equal to  $m$  is then given by

$$C_{k:l,m}^{\text{Bern}}(\mathbf{u}) = \int_{[0,1]^d} C_{k:l}(\mathbf{w}) d\mu_{m,\mathbf{u}}(\mathbf{w}), \quad \mathbf{u} \in [0, 1]^d. \quad (4.4)$$

The latter is clearly a special case of  $C_{k:l}^{\nu}$  in (4.3). If, additionally,  $m = l - k + 1$ , that is, if the smoothing distributions satisfy  $\nu_{\mathbf{u}}^{\mathcal{X}_{k:l}} = \mu_{l-k+1,\mathbf{u}}$ ,  $\mathbf{u} \in [0, 1]^d$ ,  $C_{k:l}^{\nu}$  in (4.3) or, equivalently,  $C_{k:l,m}^{\text{Bern}}$  in (4.4), corresponds to the empirical beta copula of  $\mathcal{X}_{k:l}$  studied in Segers et al. (2017).

For any  $m \in \mathbb{N}$ ,  $\mathbf{x} \in (\mathbb{R}^d)^m$ ,  $\mathbf{r} \in [0, m]^d$  and  $\mathbf{u} \in [0, 1]^d$ , let

$$\mathcal{K}_{\mathbf{r}}^{\mathbf{x}}(\mathbf{u}) = \int_{[0,1]^d} \mathbf{1}(\mathbf{r}/m \leq \mathbf{w}) d\nu_{\mathbf{u}}^{\mathbf{x}}(\mathbf{w}) = \mathbb{E} \{ \mathbf{1}(\mathbf{r}/m \leq \mathbf{W}_{\mathbf{u}}^{\mathbf{x}}) \}. \quad (4.5)$$

By linearity of the integral, we can then express  $C_{k:l}^{\nu}$  in (4.3) as

$$C_{k:l}^{\nu}(\mathbf{u}) = \frac{1}{l - k + 1} \sum_{i=k}^l \mathcal{K}_{\mathbf{R}_i^{k:l}}^{\mathcal{X}_{k:l}}(\mathbf{u}), \quad \mathbf{u} \in [0, 1]^d. \quad (4.6)$$

Since copulas have standard uniform margins, it is particularly meaningful to focus on estimators of the form (4.6) that have standard uniform margins. As verified in Section 3.1 of Kojadinovic and Yi (2022), the following two assumptions imply the latter.

**Condition 4.2.3** (No ties). With probability 1, there are no ties in each of the component samples  $X_{1j}, \dots, X_{nj}$ ,  $j \in \{1, \dots, d\}$ , of  $\mathcal{X}_{1:n}$ .

**Condition 4.2.4** (Condition for uniform margins). For any  $m \in \mathbb{N}$ ,  $\mathbf{x} \in (\mathbb{R}^d)^m$ ,  $\mathbf{u} \in [0, 1]^d$  and  $j \in \{1, \dots, d\}$ ,  $W_{j, u_j}^{\mathbf{x}}$  takes its values in the set  $\{0, 1/m, \dots, (m-1)/m, 1\}$ .

Under Condition 4.2.4, from Section 3.2 of [Kojadinovic and Yi \(2022\)](#), for any  $m \in \mathbb{N}$ ,  $\mathbf{x} \in (\mathbb{R}^d)^m$ ,  $\mathbf{r} \in [1, m]^d$  and  $\mathbf{u} \in [0, 1]^d$ ,  $\mathcal{K}_{\mathbf{r}}^{\mathbf{x}}(\mathbf{u})$  in (4.5) can be written as

$$\mathcal{K}_{\mathbf{r}}^{\mathbf{x}}(\mathbf{u}) = \bar{C}_{\mathbf{u}}^{\mathbf{x}} \left[ \bar{F}_{1, u_1}^{\mathbf{x}} \left\{ (r_1 - 1)/m \right\}, \dots, \bar{F}_{d, u_d}^{\mathbf{x}} \left\{ (r_d - 1)/m \right\} \right], \quad (4.7)$$

where  $\bar{C}_{\mathbf{u}}^{\mathbf{x}}$  (resp.  $\bar{F}_{1, u_1}^{\mathbf{x}}, \dots, \bar{F}_{d, u_d}^{\mathbf{x}}$ ) is a survival copula (resp. are the marginal survival functions) of the random vector  $\mathbf{W}_{\mathbf{u}}^{\mathbf{x}}$ . Upon additionally assuming the following two conditions considered in Section 3.2 of [Kojadinovic and Yi \(2022\)](#), estimators of the form (4.6) can be shown to be genuine copulas.

**Condition 4.2.5** (Condition on the smoothing survival margins). For any  $m \in \mathbb{N}$ ,  $\mathbf{x} \in (\mathbb{R}^d)^m$ ,  $j \in \{1, \dots, d\}$  and  $w \in [0, 1]$ , the function  $t \mapsto \bar{F}_{j, t}^{\mathbf{x}}(w)$  is right-continuous and increasing on  $[0, 1]$ .

**Condition 4.2.6** (Condition on the smoothing survival copulas). For any  $m \in \mathbb{N}$ ,  $\mathbf{x} \in (\mathbb{R}^d)^m$  and  $\mathbf{u} \in [0, 1]^d$ , the copulas  $\bar{C}_{\mathbf{u}}^{\mathbf{x}}$  in (4.7) do not depend on  $\mathbf{u}$ , that is,  $\bar{C}_{\mathbf{u}}^{\mathbf{x}} = \bar{C}^{\mathbf{x}}$ .

The following result was then proven in [Kojadinovic and Yi \(2022\)](#); see Proposition 11 and Corollary 12 therein.

**Proposition 4.2.7** ( $C_{k:l}^{\nu}$  is a genuine copula). Assume that Conditions 4.2.3, 4.2.4, 4.2.5 and 4.2.6 hold. Then, the smooth empirical copula  $C_{k:l}^{\nu}$  in (4.3) or in (4.6) can be expressed, for any  $\mathbf{u} \in [0, 1]^d$ , as

$$C_{k:l}^{\nu}(\mathbf{u}) = \frac{1}{l-k+1} \sum_{i=k}^l \bar{C}^{\mathbf{x}_{k:l}} \left\{ \bar{F}_{1, u_1}^{\mathbf{x}_{k:l}} \left( \frac{R_{i1}^{k:l} - 1}{l-k+1} \right), \dots, \bar{F}_{d, u_d}^{\mathbf{x}_{k:l}} \left( \frac{R_{id}^{k:l} - 1}{l-k+1} \right) \right\}, \quad (4.8)$$

and is a genuine copula.

From Remark 4.2.2 above, we can infer that the empirical beta copula of  $\mathbf{X}_{k:l}$  studied in [Segers et al. \(2017\)](#) is of the form (4.8) with  $\bar{C}^{\mathbf{x}_{k:l}}$  the independence copula and, for any  $j \in \{1, \dots, d\}$  and  $u \in [0, 1]$ ,  $\bar{F}_{j, u}^{\mathbf{x}_{k:l}}$  the survival function of a scaled (by  $1/(l-k+1)$ ) Binomial( $l-k+1, u$ ) random variable. For that reason, the latter will be denoted as  $C_{k:l}^{\text{Bin}}$  as we continue. As a possible improvement of the empirical beta copula  $C_{k:l}^{\text{Bin}}$  of  $\mathbf{X}_{k:l}$ , [Kojadinovic and Yi \(2022\)](#) suggested to consider a smooth data-adaptive empirical copula of the form (4.8) with  $\bar{C}^{\mathbf{x}_{k:l}}$  the empirical beta copula  $C_{k:l}^{\text{Bin}}$  and, for any  $j \in \{1, \dots, d\}$  and  $u \in [0, 1]$ ,  $\bar{F}_{j, u}^{\mathbf{x}_{k:l}}$  the survival function of a scaled (by  $1/(l-k+1)$ ) Beta-Binomial( $m, \alpha, \beta$ ) random variable, where  $m = l-k+1$ ,  $\alpha = u(m-\rho)/(\rho-1)$ ,  $\beta = (1-u)(m-\rho)/(\rho-1)$  and  $\rho = 4$ . The resulting data-adaptive estimator, denoted by  $C_{k:l}^{\text{BetaB4}}$  as we continue, was found to outperform the empirical beta copula  $C_{k:l}^{\text{Bin}}$  in terms of integrated mean squared error in all the bivariate and trivariate experiments considered in [Kojadinovic and Yi \(2022\)](#).

### 4.2.2 Asymptotics of related sequential processes

We can now define the sequential empirical processes corresponding to the empirical copula in (4.2) and to its smooth generalizations in (4.3). Let  $\Lambda = \{(s, t) \in [0, 1]^2 : s \leq t\}$  and let  $\lambda_n(s, t) = (\lfloor nt \rfloor - \lfloor ns \rfloor)/n$ ,  $(s, t) \in \Lambda$ . The corresponding two-sided sequential empirical copula processes are given, for any  $(s, t) \in \Lambda$  and  $\mathbf{u} \in [0, 1]^d$ , by

$$\mathbb{C}_n(s, t, \mathbf{u}) = \sqrt{n}\lambda_n(s, t)\{C_{\lfloor ns \rfloor+1:\lfloor nt \rfloor}(\mathbf{u}) - C(\mathbf{u})\}, \quad (4.9)$$

$$\mathbb{C}_n^\nu(s, t, \mathbf{u}) = \sqrt{n}\lambda_n(s, t)\{C_{\lfloor ns \rfloor+1:\lfloor nt \rfloor}^\nu(\mathbf{u}) - C(\mathbf{u})\}, \quad (4.10)$$

where  $C_{\lfloor ns \rfloor+1:\lfloor nt \rfloor}$  and  $C_{\lfloor ns \rfloor+1:\lfloor nt \rfloor}^\nu$  are generically defined in (4.2) and (4.3), respectively. The asymptotics of  $\mathbb{C}_n$  were established in Bücher and Kojadinovic (2016), while the asymptotics of  $\mathbb{C}_n^\nu$  (which we recall in Theorem 4.2.10 hereafter) were investigated in Kojadinovic and Yi (2022) by generalizing the arguments used in Segers et al. (2017).

The following conditions were considered in Kojadinovic and Yi (2022).

**Condition 4.2.8** (Smooth partial derivatives). For any  $j \in \{1, \dots, d\}$ , the partial derivative  $\dot{C}_j = \partial C / \partial u_j$  exists and is continuous on the set  $V_j = \{\mathbf{u} \in [0, 1]^d : u_j \in (0, 1)\}$ .

**Condition 4.2.9** (Variance condition). There exists a constant  $\kappa > 0$  such that, for any  $n \in \mathbb{N}$ ,  $\mathbf{x} \in (\mathbb{R}^d)^n$ ,  $\mathbf{u} \in [0, 1]^d$  and  $j \in \{1, \dots, d\}$ ,  $\text{Var}(W_{j, u_j}^{\mathbf{x}}) \leq \kappa u_j(1 - u_j)/n$ .

The first condition was initially considered in Segers (2012) and can be considered non-restricted as explained in the latter reference. In the rest of the paper, for any  $j \in \{1, \dots, d\}$ ,  $\dot{C}_j$  is arbitrarily defined to be zero on the set  $\{\mathbf{u} \in [0, 1]^d : u_j \in \{0, 1\}\}$ , which implies that, under Condition 4.2.8,  $\dot{C}_j$  is defined on the whole of  $[0, 1]^d$ . The second condition imposes constraints on the spread of the smoothing distributions involved in the definition of the smooth, possibly data-adaptive, empirical copulas.

**Theorem 4.2.10** (Asymptotics of  $\mathbb{C}_n^\nu$ ). *Assume that Conditions 4.2.8 and 4.2.9 hold, and that  $\mathbb{C}_n \rightsquigarrow \mathbb{C}_C$  in  $\ell^\infty(\Lambda \times [0, 1]^d)$ , where the trajectories of the limiting process  $\mathbb{C}_C$  are continuous almost surely. Then,*

$$\sup_{\substack{(s, t) \in \Lambda \\ \mathbf{u} \in [0, 1]^d}} |\mathbb{C}_n^\nu(s, t, \mathbf{u}) - \mathbb{C}_n(s, t, \mathbf{u})| = o_{\mathbb{P}}(1).$$

Consequently,  $\mathbb{C}_n^\nu \rightsquigarrow \mathbb{C}_C$  in  $\ell^\infty(\Lambda \times [0, 1]^d)$ .

Hence, the smooth sequential empirical copula process  $\mathbb{C}_n^\nu$  in (4.10) and the classical sequential empirical copula process  $\mathbb{C}_n$  in (4.9) are asymptotically equivalent when the latter converges weakly to a limiting process whose trajectories are continuous almost surely. As discussed in Section 3 of Bücher and Kojadinovic (2016), for such a convergence to hold, it suffices that the corresponding “uniform multivariate sequential empirical process” converges weakly to a limiting process whose trajectories are continuous almost surely. Specifically, let  $\mathbf{U}_1, \dots, \mathbf{U}_n$

be the unobservable sample obtained from  $\mathcal{X}_{1:n} = (\mathbf{X}_1, \dots, \mathbf{X}_n)$  by the probability integral transformations  $U_{ij} = F_j(X_{ij})$ ,  $i \in \{1, \dots, n\}$ ,  $j \in \{1, \dots, d\}$ , and let

$$\mathbb{B}_n(s, t, \mathbf{u}) = \frac{1}{\sqrt{n}} \sum_{i=[ns]+1}^{\lfloor nt \rfloor} \{\mathbf{1}(\mathbf{U}_i \leq \mathbf{u}) - C(\mathbf{u})\}, \quad (s, t, \mathbf{u}) \in \Lambda \times [0, 1]^d, \quad (4.11)$$

with the convention that  $\mathbb{B}_n(s, t, \cdot) = 0$  if  $\lfloor nt \rfloor - [ns] = 0$ . The aforementioned sufficient condition can then be stated as follows.

**Condition 4.2.11** (Weak convergence of  $\mathbb{B}_n(0, \cdot, \cdot)$ ). The sequential empirical process  $\mathbb{B}_n(0, \cdot, \cdot)$  converges weakly in  $\ell^\infty([0, 1]^{d+1})$  to a tight centered Gaussian process  $\mathbb{Z}_C$  concentrated on

$$\{f \in \mathcal{C}([0, 1]^{d+1}) : f(s, \mathbf{u}) = 0 \text{ if one of the components of } (s, \mathbf{u}) \text{ is } 0, \\ \text{and } f(s, 1, \dots, 1) = 0 \text{ for all } s \in (0, 1]\}.$$

Under Condition 4.2.11 (which holds for instance when  $(\mathbf{X}_i)_{i \in \mathbb{Z}}$  is *strongly mixing*; see, e.g., Bücher (2015) as well as forthcoming Section 4.4.1), it immediately follows from the continuous mapping theorem that  $\mathbb{B}_n \rightsquigarrow \mathbb{B}_C$  in  $\ell^\infty(\Lambda \times [0, 1]^d)$ , where

$$\mathbb{B}_C(s, t, \mathbf{u}) = \mathbb{Z}_C(t, \mathbf{u}) - \mathbb{Z}_C(s, \mathbf{u}), \quad (s, t, \mathbf{u}) \in \Lambda \times [0, 1]^d. \quad (4.12)$$

For any  $j \in \{1, \dots, d\}$  and any  $\mathbf{u} \in [0, 1]^d$ , let  $\mathbf{u}^{(j)}$  be the vector of  $[0, 1]^d$  defined by  $u_i^{(j)} = u_j$  if  $i = j$  and 1 otherwise. The following result is then an immediate consequence of Theorem 3.4 in Bücher and Kojadinovic (2016) and Proposition 3.3 of Bücher et al. (2014).

**Theorem 4.2.12** (Asymptotics of  $\mathbb{C}_n$ ). Under Conditions 4.2.8 and 4.2.11,

$$\sup_{(s, t, \mathbf{u}) \in \Lambda \times [0, 1]^d} \left| \mathbb{C}_n(s, t, \mathbf{u}) - \tilde{\mathbb{C}}_n(s, t, \mathbf{u}) \right| = o_{\mathbb{P}}(1),$$

where

$$\tilde{\mathbb{C}}_n(s, t, \mathbf{u}) = \mathbb{B}_n(s, t, \mathbf{u}) - \sum_{j=1}^d \dot{C}_j(\mathbf{u}) \mathbb{B}_n(s, t, \mathbf{u}^{(j)}), \quad (s, t, \mathbf{u}) \in \Lambda \times [0, 1]^d, \quad (4.13)$$

and  $\mathbb{B}_n$  is defined in (4.11). Consequently,  $\mathbb{C}_n \rightsquigarrow \mathbb{C}_C$  in  $\ell^\infty(\Lambda \times [0, 1]^d)$ , where

$$\mathbb{C}_C(s, t, \mathbf{u}) = \mathbb{B}_C(s, t, \mathbf{u}) - \sum_{j=1}^d \dot{C}_j(\mathbf{u}) \mathbb{B}_C(s, t, \mathbf{u}^{(j)}), \quad (s, t, \mathbf{u}) \in \Lambda \times [0, 1]^d, \quad (4.14)$$

and  $\mathbb{B}_C$  is defined in (4.12).

We end this section with the statement of a corollary of Theorems 4.2.10 and 4.2.12. Having (4.5) in mind, two natural smooth extensions of the unobservable empirical process  $\mathbb{B}_n$  in (4.11) can

be defined, for any  $(s, t, \mathbf{u}) \in \Lambda \times [0, 1]^d$ , by

$$\tilde{\mathbb{B}}_n^\nu(s, t, \mathbf{u}) = \frac{1}{\sqrt{n}} \sum_{i=[ns]+1}^{[nt]} \left\{ \int_{[0,1]^d} \mathbf{1}(U_i \leq \mathbf{w}) d\nu_{\mathbf{u}^{[ns]+1:[nt]}}(\mathbf{w}) - C(\mathbf{u}) \right\}, \quad (4.15)$$

$$\bar{\mathbb{B}}_n^\nu(s, t, \mathbf{u}) = \frac{1}{\sqrt{n}} \sum_{i=[ns]+1}^{[nt]} \left\{ \int_{[0,1]^d} \mathbf{1}(U_i \leq \mathbf{w}) d\nu_{\mathbf{u}^{1:n}}(\mathbf{w}) - C(\mathbf{u}) \right\}. \quad (4.16)$$

Combining Theorem 4.2.12 with key intermediate results used in Kojadinovic and Yi (2022) for proving Theorem 4.2.10 stated above, we obtain the following asymptotic representations for the smooth sequential empirical process  $\mathbb{C}_n^\nu$  in (4.10). The proof of this result is given in Appendix 4.7.1.

**Corollary 4.2.13** (Asymptotic representations of  $\mathbb{C}_n^\nu$ ). *Under Conditions 4.2.8, 4.2.9 and 4.2.11,*

$$\begin{aligned} \sup_{(s,t,\mathbf{u}) \in \Lambda \times [0,1]^d} \left| \mathbb{C}_n^\nu(s, t, \mathbf{u}) - \tilde{\mathbb{C}}_n^\nu(s, t, \mathbf{u}) \right| &= o_{\mathbb{P}}(1), \\ \sup_{(s,t,\mathbf{u}) \in \Lambda \times [0,1]^d} \left| \mathbb{C}_n^\nu(s, t, \mathbf{u}) - \bar{\mathbb{C}}_n^\nu(s, t, \mathbf{u}) \right| &= o_{\mathbb{P}}(1), \end{aligned}$$

where, for any  $(s, t, \mathbf{u}) \in \Lambda \times [0, 1]^d$ ,

$$\begin{aligned} \tilde{\mathbb{C}}_n^\nu(s, t, \mathbf{u}) &= \tilde{\mathbb{B}}_n^\nu(s, t, \mathbf{u}) - \sum_{j=1}^d \dot{C}_j(\mathbf{u}) \tilde{\mathbb{B}}_n^\nu(s, t, \mathbf{u}^{(j)}), \\ \bar{\mathbb{C}}_n^\nu(s, t, \mathbf{u}) &= \bar{\mathbb{B}}_n^\nu(s, t, \mathbf{u}) - \sum_{j=1}^d \dot{C}_j(\mathbf{u}) \bar{\mathbb{B}}_n^\nu(s, t, \mathbf{u}^{(j)}). \end{aligned}$$

*Remark 4.2.14.* The previous results do not unfortunately allow us to decide which of the above two asymptotic representations for  $\mathbb{C}_n^\nu$  may be better. The knowledge of the underlying convergence rates would be needed for that. As we shall see in Section 4.4, these representations will be at the root of smooth proposals for bootstrapping  $\mathbb{C}_n^\nu$  in a time series context.

### 4.3 Bootstrap by drawing samples from the estimators in the i.i.d. case

The aim of this section is to study both theoretically and empirically a smooth bootstrap *à la* Kiriliouk et al. (2021) based on drawing samples from the smooth estimators defined in the previous section. As hinted at in the introduction, such an approach can only be used in the i.i.d. case. Throughout this section, we thus assume that the random vectors in  $\mathcal{X}_{1:n}$  are i.i.d. Notice that the latter implies Condition 4.2.3. Given that change-point analysis is essentially of interest in the time series setting, we do not consider a sequential setting below but instead focus only on the situation where  $k = 1$  and  $l = n$ .

This section is organized as follows. After describing the sampling algorithm on which the smooth bootstrap is based, we state conditions under which it is asymptotically valid and report results of finite-sample experiments comparing smooth bootstraps based on the empirical beta copula  $C_{1:n}^{\text{Bin}}$  and on its data-adaptive extension  $C_{1:n}^{\text{BetaB4}}$  proposed in [Kojadinovic and Yi \(2022\)](#) and recalled at the end of [Section 4.2.1](#).

### 4.3.1 Drawing samples from the smooth empirical copulas

As explained in [Section 4.2.1](#), the empirical beta copula  $C_{1:n}^{\text{Bin}}$  is a particular case of the smooth estimators  $C_{1:n}^\nu$  defined in [\(4.3\)](#). From [Segers et al. \(2017\)](#) (see also Lemma 1 in [Kojadinovic and Yi 2022](#)), one has that

$$C_{1:n}^{\text{Bin}}(\mathbf{u}) = \frac{1}{n} \sum_{i=1}^n \prod_{j=1}^d F_{n, R_{ij}^{1:n}}(u_j), \quad \mathbf{u} = (u_1, \dots, u_d) \in [0, 1]^d, \quad (4.17)$$

where, for any  $n \in \mathbb{N}$  and  $r \in \{1, \dots, n\}$ ,  $F_{n,r}$  denotes the d.f. of the beta distribution with shape parameters  $\alpha = r$  and  $\beta = n + 1 - r$ . It follows from [\(4.17\)](#) that  $C_{1:n}^{\text{Bin}}$  is a mixture of  $n$   $d$ -dimensional distributions having beta margins and whose copula is the independence copula. To generate one random variate from  $C_{1:n}^{\text{Bin}}$ , it thus suffices to randomly select one of the  $n$  components of the mixture by drawing a uniform on  $\{1, \dots, n\}$  and then generate one random variate from the selected  $d$ -dimensional distribution. This is detailed in [Algorithm 3.2](#) of [Kiriliouk et al. \(2021\)](#).

In a related way, having [\(4.6\)](#) in mind, it thus suffices to assume the following to be able sample from  $C_{1:n}^\nu$ .

**Condition 4.3.1.** ( $C_{1:n}^\nu$  is a mixture) For any  $n \in \mathbb{N}$ ,  $\mathbf{x} \in (\mathbb{R}^d)^n$  and  $\mathbf{r} \in \{1, \dots, n\}^d$ ,  $\mathcal{K}_{\mathbf{r}}^{\mathbf{x}}$  in [\(4.5\)](#) is a d.f. on  $[0, 1]^d$ .

The sampling algorithm is then conceptually the same as [Algorithm 3.2](#) of [Kiriliouk et al. \(2021\)](#).

**Algorithm 4.3.2.** (*Sampling from  $C_{1:n}^\nu$  under Condition 4.3.1*)

1. Generate  $I$  from the discrete uniform distribution on  $\{1, \dots, n\}$ .
2. Generate a random variate  $\mathbf{V}^\#$  from a  $d$ -dimensional distribution whose d.f. is  $\mathcal{K}_{\mathbf{r}_I}^{\mathbf{x}_{1:n}}$ .

The above algorithm can be used in practice as soon as one knows how to sample from the d.f.s  $\mathcal{K}_{\mathbf{r}_i}^{\mathbf{x}_{1:n}}$ ,  $i \in \{1, \dots, n\}$ .

Interestingly enough, three of the conditions stated in [Section 4.2.1](#) imply [Condition 4.3.1](#) as shown in the next result proven in [Appendix 4.7.2](#).

**Proposition 4.3.3.** *Conditions 4.2.4, 4.2.5 and 4.2.6 imply Condition 4.3.1. Specifically, under Conditions 4.2.4, 4.2.5 and 4.2.6, for any  $n \in \mathbb{N}$ ,  $\mathbf{x} \in (\mathbb{R}^d)^n$  and  $\mathbf{r} \in \{1, \dots, n\}^d$ ,  $\mathcal{K}_{\mathbf{r}}^{\mathbf{x}}$  in (4.5) is a d.f. on  $[0, 1]^d$  whose  $d$  univariate margins, denoted by  $\mathcal{K}_{r_1, 1}^{\mathbf{x}}, \dots, \mathcal{K}_{r_d, d}^{\mathbf{x}}$ , respectively, satisfy  $\mathcal{K}_{r_j, j}^{\mathbf{x}}(u) = \bar{\mathcal{F}}_{j, u}^{\mathbf{x}}\{(r_j - 1)/n\}$ ,  $u \in [0, 1]$ ,  $j \in \{1, \dots, d\}$ , and whose copula is  $\bar{\mathcal{C}}^{\mathbf{x}}$ .*

*Remark 4.3.4.* The previous result leads to an alternative (and simpler) proof of Proposition 11 of [Kojadinovic and Yi \(2022\)](#). Indeed, under the assumptions of Proposition 4.3.3,  $C_{1:n}^{\nu}$  in (4.6) is a convex combination of multivariate d.f.s on  $[0, 1]^d$  and therefore a multivariate d.f. on  $[0, 1]^d$ . Since Condition 4.2.3 holds in the current i.i.d. setting, from Section 3.1 in [Kojadinovic and Yi \(2022\)](#), Condition 4.2.4 also implies that  $C_{1:n}^{\nu}$  has standard uniform margins. Hence, under the assumptions of Proposition 4.3.3,  $C_{1:n}^{\nu}$  is a genuine copula.

For any univariate d.f.  $H$ , let  $H^{-1}$  denote its associated quantile function (generalized inverse) defined by  $H^{-1}(y) = \inf\{x \in \mathbb{R} : H(x) \geq y\}$ ,  $y \in [0, 1]$ , with the convention that  $\inf \emptyset = \infty$ . The second step of Algorithm 4.3.2 can then be made explicit under Conditions 4.2.4, 4.2.5 and 4.2.6 :

- (i) Generate a random variate  $\mathbf{U}^{\#}$  from the copula  $\bar{\mathcal{C}}^{\mathbf{x}_{1:n}}$  independently of  $I$ .
- (ii) A random variate from the distribution whose d.f. is  $\mathcal{K}_{\mathbf{R}_{1:n}^{\mathbf{x}_{1:n}}}$  is then

$$\mathbf{V}^{\#} = \left( \mathcal{K}_{R_{11}^{1:n}, 1}^{\mathbf{x}_{1:n}, -1}(\mathbf{U}_1^{\#}), \dots, \mathcal{K}_{R_{1d}^{1:n}, d}^{\mathbf{x}_{1:n}, -1}(\mathbf{U}_d^{\#}) \right). \quad (4.18)$$

We end this section by discussing how Algorithm 4.3.2 can be practically implemented for the smooth data-adaptive estimator  $C_{1:n}^{\text{BetaB4}}$  introduced in [Kojadinovic and Yi \(2022\)](#) as a possible improvement of the empirical beta copula  $C_{1:n}^{\text{Bin}}$ . Recall from Section 4.2.1 that  $C_{1:n}^{\text{BetaB4}}$  is of the form (4.8) with  $\bar{\mathcal{C}}^{\mathbf{x}_{1:n}}$  the empirical beta copula  $C_{1:n}^{\text{Bin}}$  and, for any  $j \in \{1, \dots, d\}$  and  $u \in [0, 1]$ ,  $\bar{\mathcal{F}}_{j, u}^{\mathbf{x}_{1:n}}$  the survival function of a scaled (by  $1/n$ ) Beta-Binomial( $n, \alpha, \beta$ ) random variable, where  $\alpha = u(n - \rho)/(\rho - 1)$ ,  $\beta = (1 - u)(n - \rho)/(\rho - 1)$  and  $\rho = 4$ . The latter implies that, for any  $i \in \{1, \dots, n\}$ ,  $j \in \{1, \dots, d\}$  and  $u \in [0, 1]$ ,

$$\mathcal{K}_{R_{ij}^{1:n}, j}^{\mathbf{x}_{1:n}}(u) = \bar{\mathcal{F}}_{j, u}^{\mathbf{x}_{1:n}}\{(R_{ij}^{1:n} - 1)/n\} = \mathbb{P}(nW_{j, u}^{\mathbf{x}_{1:n}} > R_{ij}^{1:n} - 1) = \bar{\mathcal{B}}_{n, u, \rho}(R_{ij}^{1:n} - 1), \quad (4.19)$$

where  $\bar{\mathcal{B}}_{n, u, \rho}$  is the survival function of the Beta-Binomial( $n, \alpha, \beta$ ). As can be checked from Lemma 27 in [Kojadinovic and Yi \(2022\)](#) and Lemma 4.7.4 in Appendix 4.7.2, the univariate d.f.  $\mathcal{K}_{R_{ij}^{1:n}, j}^{\mathbf{x}_{1:n}}$  in (4.19) is continuous and strictly increasing, respectively. Hence, to compute its associated quantile function needed in (4.18), one can proceed numerically. In that respect, an implementation of Algorithm 4.3.2 for the R statistical environment for the estimators  $C_{1:n}^{\text{Bin}}$  and  $C_{1:n}^{\text{BetaB4}}$  is available on the web page of the first author.

### 4.3.2 Asymptotic validity results

Building upon the work of Kiriliouk et al. (2021), we will now provide asymptotic validity results for a smooth bootstrap based on drawing samples from  $C_{1:n}^\nu$  in (4.6) under Conditions 4.2.4, 4.2.5 and 4.2.6. Recall that, according to Proposition 4.3.3, the latter conditions imply Condition 4.3.1. Let  $\mathbf{V}_{1:n}^\# = (\mathbf{V}_1^\#, \dots, \mathbf{V}_n^\#)$  be a random sample from  $C_{1:n}^\nu$  obtained by applying Algorithm 4.3.2  $n$  times independently. Note that this implies that the component samples of  $\mathbf{V}_{1:n}^\#$  do not contain ties with probability 1. For any  $j \in \{1, \dots, d\}$ , let  $G_{1:n,j}^\#$  be the empirical d.f. computed from the  $j$ th component sample  $V_{1j}^\#, \dots, V_{nj}^\#$  of  $\mathbf{V}_{1:n}^\#$ . Then,  $R_{ij}^{1:n,\#} = nG_{1:n,j}^\#(V_{ij}^\#)$  is the rank of  $V_{ij}^\#$  among  $V_{1j}^\#, \dots, V_{nj}^\#$ . The (classical) empirical copula of  $\mathbf{V}_{1:n}^\#$  is thus given by

$$C_{1:n}^\#(\mathbf{u}) = \frac{1}{n} \sum_{i=1}^n \prod_{j=1}^d \mathbf{1} \left( \frac{R_{ij}^{1:n,\#}}{n} \leq u_j \right), \quad \mathbf{u} \in [0, 1]^d, \quad (4.20)$$

and the smooth analog of  $C_{1:n}^\nu$  for  $\mathbf{V}_{1:n}^\#$  is

$$C_{1:n}^{\#,\nu}(\mathbf{u}) = \int_{[0,1]^d} C_{1:n}^\#(\mathbf{w}) d\nu_{\mathbf{u}^{\mathbf{V}_{1:n}^\#}}(\mathbf{w}), \quad \mathbf{u} \in [0, 1]^d. \quad (4.21)$$

To state our asymptotic validity results, we consider independent copies  $\mathbf{V}_{1:n}^{\#[1]}, \mathbf{V}_{1:n}^{\#[2]}, \dots$  of  $\mathbf{V}_{1:n}^\#$ . Let  $C_{1:n}^{\#[i]}$  (resp.  $C_{1:n}^{\#,\nu,[i]}$ ) be the version of  $C_{1:n}^\#$  in (4.20) (resp.  $C_{1:n}^{\#,\nu}$  in (4.21)) obtained from  $\mathbf{V}_{1:n}^{\#[i]}$ ,  $i \in \mathbb{N}$ .

The following result can be regarded as an extension of Proposition 3.3 of Kiriliouk et al. (2021) and is proven in Appendix 4.7.3.

**Theorem 4.3.5.** *Assume that the random vectors in  $\mathcal{X}_{1:n}$  are i.i.d., and that Conditions 4.2.4, 4.2.5, 4.2.6, 4.2.8 and 4.2.9 hold. Then,*

$$\begin{aligned} & (\mathbb{C}_n(0, 1, \cdot), \sqrt{n}(C_{1:n}^{\#[1]} - C_{1:n}), \sqrt{n}(C_{1:n}^{\#[2]} - C_{1:n})) \\ & \rightsquigarrow (\mathbb{C}_C(0, 1, \cdot), \mathbb{C}_C^{[1]}(0, 1, \cdot), \mathbb{C}_C^{[2]}(0, 1, \cdot)), \\ & (\mathbb{C}_n^\nu(0, 1, \cdot), \sqrt{n}(C_{1:n}^{\#,\nu,[1]} - C_{1:n}^\nu), \sqrt{n}(C_{1:n}^{\#,\nu,[2]} - C_{1:n}^\nu)) \\ & \rightsquigarrow (\mathbb{C}_C(0, 1, \cdot), \mathbb{C}_C^{[1]}(0, 1, \cdot), \mathbb{C}_C^{[2]}(0, 1, \cdot)) \end{aligned}$$

in  $\{\ell^\infty([0, 1]^d)\}^3$ , where  $\mathbb{C}_n$  and  $\mathbb{C}_n^\nu$  are defined in (4.9) and (4.10), respectively, and  $\mathbb{C}_C^{[1]}$  and  $\mathbb{C}_C^{[2]}$  are independent copies of  $\mathbb{C}_C$  defined in (4.14).

*Remark 4.3.6.* The first joint weak convergence in Theorem 4.3.5 establishes the asymptotic validity of a smooth bootstrap for the (non-sequential) classical empirical process while the second one provides a similar results for the smooth empirical copula process  $\mathbb{C}_n^\nu(0, 1, \cdot)$ . According to Lemma 3.1 in Bücher and Kojadinovic (2019), these two joint weak convergences are equivalent to similar joint weak convergences with  $B \geq 2$  bootstrap replicates. In a further step, the latter can be transferred to the “statistic level” using the continuous mapping theorem



TABLE 4.1: Coverage probabilities (cov.) and average lengths (ave.) of 95%-confidence intervals for Kendall’s tau estimated from 1000 random samples of size  $n \in \{20, 40, 80, 160\}$  from the bivariate Clayton or Gumbel–Hougaard copula with a Kendall’s tau of  $\tau \in \{0, 0.5, 0.75, 0.9\}$ . Each confidence interval was computed using 1000 smooth bootstrap samples drawn from either  $C_{1:n}^{\text{Bin}}$  or  $C_{1:n}^{\text{BetaB4}}$  using Algorithm 4.3.2.

		Clayton				Gumbel–Hougaard			
		Bin		BetaB4		Bin		BetaB4	
$\tau$	$n$	cov.	ave.	cov.	ave.	cov.	ave.	cov.	ave.
0.00	20	0.973	0.626	0.967	0.615	0.969	0.624	0.962	0.614
	40	0.962	0.428	0.954	0.424	0.950	0.430	0.944	0.425
	80	0.949	0.298	0.943	0.296	0.970	0.297	0.961	0.296
	160	0.944	0.208	0.943	0.207	0.949	0.208	0.949	0.208
0.50	20	0.971	0.513	0.972	0.493	0.978	0.521	0.982	0.498
	40	0.958	0.347	0.954	0.334	0.959	0.345	0.957	0.332
	80	0.946	0.239	0.938	0.233	0.950	0.237	0.947	0.231
	160	0.954	0.168	0.957	0.165	0.954	0.164	0.958	0.162
0.75	20	0.717	0.392	0.899	0.367	0.777	0.391	0.927	0.358
	40	0.728	0.234	0.908	0.221	0.793	0.231	0.954	0.211
	80	0.798	0.151	0.930	0.146	0.844	0.146	0.953	0.137
	160	0.866	0.101	0.943	0.100	0.883	0.098	0.944	0.094
0.90	20	0.000	0.315	0.212	0.272	0.000	0.317	0.270	0.264
	40	0.000	0.160	0.475	0.131	0.000	0.162	0.593	0.127
	80	0.000	0.086	0.692	0.074	0.000	0.087	0.804	0.069
	160	0.000	0.050	0.837	0.047	0.000	0.050	0.902	0.043

or the functional delta method, which could then be combined with the results in Section 4 of [Bücher and Kojadinovic \(2019\)](#) to establish the validity of bootstrap-based confidence intervals or tests. Note also that, from Lemma 3.1 in [Bücher and Kojadinovic \(2019\)](#), the unconditional asymptotic validity results appearing in Theorem 4.3.5 are equivalent to possibly more classical conditional results which rely, however, on a more subtle mode of convergence. For instance, the first claim can be equivalently informally stated as “ $\sqrt{n}(C_{1:n}^{\#, [1]} - C_{1:n})$  converges weakly to  $\mathbb{C}_C(0, 1, \cdot)$  in  $\ell^\infty([0, 1]^d)$  conditionally on the data in probability”; see, e.g., [Kosorok \(2008, Section 2.2.3\)](#) or [Appendix 4.7.3](#) for a precise definition of that mode of convergence.

### 4.3.3 Finite-sample comparison of two smooth bootstraps

As already mentioned in the introduction, in their Monte Carlo experiments, [Kiriliouk et al. \(2021\)](#) found the smooth bootstrap based on the empirical beta copula  $C_{1:n}^{\text{Bin}}$  to be a competitive alternative to many other resampling schemes (including the multiplier bootstrap to be studied in the forthcoming section). Since the data-adaptive empirical copula  $C_{1:n}^{\text{BetaB4}}$  was found to outperform the empirical beta copula  $C_{1:n}^{\text{Bin}}$  in the experiments reported in [Kojadinovic and Yi \(2022\)](#), it seems natural to empirically investigate how the smooth bootstrap based on  $C_{1:n}^{\text{BetaB4}}$  compares to the smooth bootstrap based on  $C_{1:n}^{\text{Bin}}$ . To do so, we reproduced some of the experiments reported in Sections 4.2 and 4.3 of [Kiriliouk et al. \(2021\)](#).

TABLE 4.2: Coverage probabilities (cov.) and average lengths (ave.) of 95%-confidence intervals for the parameter of a bivariate Frank copula estimated by maximum pseudo-likelihood from 1000 random samples of size  $n \in \{20, 40, 80\}$  from the bivariate Frank copula with a Kendall's tau of  $\tau \in \{-0.9, -0.75, -0.5, 0, 0.5, 0.75, 0.9\}$ . Each confidence interval was computed using 1000 smooth bootstrap samples drawn from either  $C_{1:n}^{\text{Bin}}$  or  $C_{1:n}^{\text{BetaB4}}$  using Algorithm 4.3.2.

$\tau$	$n$	Bin		BetaB4	
		cov.	ave.	cov.	ave.
-0.90	20	0.000	0.194	0.000	0.133
	40	0.000	0.086	0.056	0.051
	80	0.000	0.039	0.313	0.023
-0.75	20	0.731	0.286	0.940	0.237
	40	0.641	0.153	0.939	0.126
	80	0.662	0.088	0.947	0.077
-0.50	20	0.988	0.548	0.981	0.511
	40	0.975	0.342	0.957	0.327
	80	0.957	0.230	0.940	0.224
0.00	20	0.952	1.009	0.946	1.002
	40	0.937	0.681	0.929	0.681
	80	0.941	0.467	0.938	0.469
0.50	20	0.986	0.542	0.970	0.508
	40	0.972	0.344	0.949	0.328
	80	0.959	0.224	0.948	0.219
0.75	20	0.722	0.285	0.938	0.235
	40	0.634	0.154	0.927	0.128
	80	0.671	0.088	0.942	0.077
0.90	20	0.000	0.193	0.000	0.132
	40	0.000	0.086	0.046	0.051
	80	0.000	0.039	0.319	0.023

We first estimated coverage probabilities and average lengths of confidence intervals of level 95% for Kendall's tau from 1000 random samples of size  $n \in \{20, 40, 80, 160\}$  from the bivariate Clayton or Gumbel–Hougaard copula with a Kendall's tau of  $\tau \in \{0, 0.5, 0.75, 0.9\}$ . Each confidence interval was computed using 1000 smooth bootstrap samples drawn from either  $C_{1:n}^{\text{Bin}}$  or  $C_{1:n}^{\text{BetaB4}}$ . The results are reported in Table 4.1. As one can see, under independence or moderate dependence ( $\tau \in \{0, 0.5\}$ ), the estimated coverage probabilities are overall on target and very similar for the two resampling schemes. The intervals obtained using the smooth bootstrap based on  $C_{1:n}^{\text{BetaB4}}$  seem nonetheless to be slightly shorter on average. Under strong dependence ( $\tau = 0.75$ ) however, the estimated coverage probabilities of the confidence intervals computed using the smooth bootstrap based on  $C_{1:n}^{\text{Bin}}$  are substantially below the 0.95 target value. The results for  $\tau = 0.9$  actually show that the smooth bootstrap based on  $C_{1:n}^{\text{Bin}}$  is unable to generate samples with such a very strong dependence. While its results are not perfect, the smooth bootstrap based on  $C_{1:n}^{\text{BetaB4}}$  copes much better with strong dependence. This is likely to be due to the modification of the “shape” of the underlying smoothing distributions using the empirical beta copula in the expression of  $C_{1:n}^{\text{BetaB4}}$  as can be deduced from (4.8).

In a second experiment, following Kiriliouk et al. (2021, Section 4.3) we estimated coverage probabilities and average lengths of 95%-confidence intervals for the parameter of a bivariate Frank copula estimated by maximum pseudo-likelihood (see Genest et al., 1995) from 1000

random samples of size  $n \in \{20, 40, 80, 160\}$  from the bivariate Frank copula with a Kendall's tau of  $\tau \in \{-0.9, -0.75, -0.5, 0, 0.5, 0.75, 0.9\}$ . Again, each confidence interval was computed using 1000 smooth bootstrap samples drawn from either  $C_{1:n}^{\text{Bin}}$  or  $C_{1:n}^{\text{BetaB4}}$ . The results are reported in Table 4.2 and the main conclusion is qualitatively the same as for the previous experiment: the smooth bootstrap based on  $C_{1:n}^{\text{BetaB4}}$  copes much better with strong dependence than the smooth bootstrap based on  $C_{1:n}^{\text{Bin}}$ .

## 4.4 Smooth sequential dependent multiplier bootstraps in the time series case

The smooth bootstrap investigated in the previous section can only be used in the case of i.i.d. observations. Fortunately, the multiplier bootstrap, one of the most popular approaches for bootstrapping functionals of the classical empirical copula, can be employed in the time series setting. In this section, after providing some intuitions and defining *multiplier sequences*, we recall the non-smooth sequential dependent multiplier bootstrap studied in Bücher and Kojadinovic (2016). We next propose smooth extensions of the latter, provide asymptotic validity results and compare the finite-sample performance of three (smooth) multiplier bootstraps for approximating three (smooth) empirical copula processes. Finally, as an application, we consider a smooth version (based on the empirical beta copula and corresponding smooth multiplier bootstrap replicates) of the test for change-point detection developed in Bücher et al. (2014) and we compare its finite-sample performance to that of its non-smooth counterpart.

### 4.4.1 Main intuition and existing work

As mentioned in Section 4.2.2, Condition 4.2.11 holds under *strong mixing*. Given a stationary time series  $(\mathbf{Y}_i)_{i \in \mathbb{Z}}$ , denote by  $\mathcal{F}_j^k$  the  $\sigma$ -field generated by  $(\mathbf{Y}_i)_{j \leq i \leq k}$ ,  $j, k \in \mathbb{Z} \cup \{-\infty, +\infty\}$ , and recall that the strong mixing coefficients corresponding to the stationary sequence  $(\mathbf{Y}_i)_{i \in \mathbb{Z}}$  are then defined by

$$\alpha_r^{\mathbf{Y}} = \sup_{A \in \mathcal{F}_{-\infty}^0, B \in \mathcal{F}_r^{+\infty}} |\mathbb{P}(A \cap B) - \mathbb{P}(A)\mathbb{P}(B)|, \quad r \in \mathbb{N}, r > 0,$$

and that the sequence  $(\mathbf{Y}_i)_{i \in \mathbb{Z}}$  is said to be *strongly mixing* if  $\alpha_r^{\mathbf{Y}} \rightarrow 0$  as  $r \rightarrow \infty$ .

From Bücher (2015), Condition 4.2.11 holds if the strong mixing coefficients of the time series  $(\mathbf{X}_i)_{i \in \mathbb{Z}}$  satisfy  $\alpha_r^{\mathbf{X}} = O(r^{-a})$  with  $a > 1$  as  $r \rightarrow \infty$ . In that case, Theorem 4.2.12 suggests that, in order to bootstrap the classical sequential empirical copula process  $\mathbb{C}_n$  in (4.9) in an asymptotically valid way, it suffices to bootstrap the process  $\tilde{\mathbb{C}}_n$  in (4.13). The latter could be done by bootstrapping  $\mathbb{B}_n$  in (4.11) and estimating the first-order partial derivatives  $\dot{C}_j$ ,  $j \in \{1, \dots, d\}$ , of  $C$ . Such an approach was initially proposed in the independent non-sequential setting by Scaillet (2005) and Rémillard and Scaillet (2009) who used a *multiplier bootstrap*

in the spirit of [van der Vaart and Wellner \(2000, Chapter 2.9\)](#) to resample  $\mathbb{B}_n$ , and finite-differencing to estimate the partial derivatives  $\dot{C}_j$ ,  $j \in \{1, \dots, d\}$ . This resampling scheme was extended to the time series sequential setting in [Bücher and Kojadinovic \(2016\)](#) and [Bücher et al. \(2014\)](#).

#### 4.4.2 I.i.d. and dependent multiplier sequences

In the case of independent observations, multiplier bootstraps are based on *i.i.d. multiplier sequences*. We say that a sequence of random variables  $(\xi_{i,n})_{i \in \mathbb{Z}}$  is an i.i.d. multiplier sequence if:

- (M0)  $(\xi_{i,n})_{i \in \mathbb{Z}}$  is i.i.d., independent of  $\mathcal{X}_{1:n}$ , with distribution not changing with  $n$ , having mean 0, variance 1, and being such that  $\int_0^\infty \{\mathbb{P}(|\xi_{0,n}| > x)\}^{1/2} dx < \infty$ .

The time series extension of the multiplier bootstrap relies on the notion of *dependent multiplier sequence*. The key idea due to [Bühlmann \(1993\)](#) is to replace i.i.d. multipliers by suitably serially dependent multipliers that will capture the serial dependence in the data. We say that a sequence of random variables  $(\xi_{i,n})_{i \in \mathbb{Z}}$  is a dependent multiplier sequence if:

- (M1) The sequence of random variables  $(\xi_{i,n})_{i \in \mathbb{Z}}$  is stationary with  $\mathbb{E}(\xi_{0,n}) = 0$ ,  $\mathbb{E}(\xi_{0,n}^2) = 1$  and  $\sup_{n \geq 1} \mathbb{E}(|\xi_{0,n}|^\gamma) < \infty$  for all  $\gamma \geq 1$ , and is independent of the available sample  $\mathcal{X}_{1:n}$ .
- (M2) There exists a sequence  $\ell_n \rightarrow \infty$  of strictly positive constants such that  $\ell_n = o(n)$  and the sequence  $(\xi_{i,n})_{i \in \mathbb{Z}}$  is  $\ell_n$ -dependent, i.e.,  $\xi_{i,n}$  is independent of  $\xi_{i+h,n}$  for all  $h > \ell_n$  and  $i \in \mathbb{N}$ .
- (M3) There exists a function  $\varphi : \mathbb{R} \rightarrow [0, 1]$ , symmetric around 0, continuous at 0, satisfying  $\varphi(0) = 1$  and  $\varphi(x) = 0$  for all  $|x| > 1$  such that  $\mathbb{E}(\xi_{0,n}\xi_{h,n}) = \varphi(h/\ell_n)$  for all  $h \in \mathbb{Z}$ .

As shall become clearer for instance from [\(4.22\)](#) or [\(4.23\)](#) below, the bandwidth parameter  $\ell_n$  defined in (M2) plays a role similar to that of the block length in the block bootstrap. In practice, for the non-smooth sequential dependent multiplier bootstrap to be presented in the forthcoming section, its value can be chosen in a data-driven way using the approach described in detail in [Bücher and Kojadinovic \(2016, Section 5\)](#); see also [Section 4.4.6](#). The latter reference also describes in detail ways to generate dependent multiplier sequences.

#### 4.4.3 Non-smooth sequential dependent multiplier replicates

Let  $(\xi_{i,n}^{[1]})_{i \in \mathbb{Z}}$ ,  $(\xi_{i,n}^{[2]})_{i \in \mathbb{Z}}, \dots$ , be independent copies of the same multiplier sequence. Two different multiplier bootstrap replicates of the process  $\mathbb{B}_n$  in [\(4.11\)](#) were proposed in [Bücher and Kojadinovic \(2016\)](#) and [Bücher et al. \(2014\)](#), respectively. For any  $b \in \mathbb{N}$ ,  $(s, t) \in \Lambda$  and  $\mathbf{u} \in [0, 1]^d$ ,

they are defined by

$$\hat{\mathbb{B}}_n^{[b]}(s, t, \mathbf{u}) = \frac{1}{\sqrt{n}} \sum_{i=[ns]+1}^{[nt]} \xi_{i,n}^{[b]} \left\{ \mathbf{1}(\hat{U}_i^{1:n} \leq \mathbf{u}) - C_{1:n}(\mathbf{u}) \right\} \quad (4.22)$$

and

$$\check{\mathbb{B}}_n^{[b]}(s, t, \mathbf{u}) = \frac{1}{\sqrt{n}} \sum_{i=[ns]+1}^{[nt]} \xi_{i,n}^{[b]} \left\{ \mathbf{1}(\hat{U}_i^{[ns]+1:[nt]} \leq \mathbf{u}) - C_{[ns]+1:[nt]}(\mathbf{u}) \right\}, \quad (4.23)$$

respectively, where  $C_{1:n}$  and  $C_{[ns]+1:[nt]}$  are generically defined in (4.2) and with the convention that  $\hat{\mathbb{B}}_n^{[b]}(s, t, \cdot) = \check{\mathbb{B}}_n^{[b]}(s, t, \cdot) = 0$  if  $[nt] - [ns] = 0$ .

In order to define multiplier bootstrap replicates of  $\tilde{\mathbb{C}}_n$  in (4.13), it is further necessary to estimate the unknown first-order partial derivatives  $\dot{C}_j$ ,  $j \in \{1, \dots, d\}$ , of  $C$ . In the rest of this section,  $\dot{C}_{j,k:l}$  will denote an estimator of  $\dot{C}_j$  based on a stretch  $\mathbf{X}_{k:l} = (\mathbf{X}_k, \dots, \mathbf{X}_l)$  of observations,  $1 \leq k \leq l \leq n$ , with the convention that  $\dot{C}_{j,k:l} = 0$  if  $k > l$ . Then, following Bücher and Kojadinovic (2016) and Bücher et al. (2014), we consider two types of multiplier bootstrap replicates of  $\mathbb{C}_n$  in (4.9). For any  $b \in \mathbb{N}$ ,  $(s, t) \in \Lambda$  and  $\mathbf{u} \in [0, 1]^d$ , these are defined by

$$\hat{\mathbb{C}}_n^{[b]}(s, t, \mathbf{u}) = \hat{\mathbb{B}}_n^{[b]}(s, t, \mathbf{u}) - \sum_{j=1}^d \dot{C}_{j,1:n}(\mathbf{u}) \hat{\mathbb{B}}_n^{[b]}(s, t, \mathbf{u}^{(j)}) \quad (4.24)$$

and

$$\check{\mathbb{C}}_n^{[b]}(s, t, \mathbf{u}) = \check{\mathbb{B}}_n^{[b]}(s, t, \mathbf{u}) - \sum_{j=1}^d \dot{C}_{j,[ns]+1:[nt]}(\mathbf{u}) \check{\mathbb{B}}_n^{[b]}(s, t, \mathbf{u}^{(j)}), \quad (4.25)$$

respectively, where  $\hat{\mathbb{B}}_n^{[b]}$  (resp.  $\check{\mathbb{B}}_n^{[b]}$ ) is defined in (4.22) (resp. (4.23)). Clearly, both types of replicates coincide in a non-sequential setting as  $\hat{\mathbb{C}}_n^{[b]}(0, 1, \cdot) = \check{\mathbb{C}}_n^{[b]}(0, 1, \cdot)$ . As far as the estimators of the partial derivatives are concerned, it is expected that the more accurate they are, the better the approximation of the “sampling distribution” of  $\mathbb{C}_n$  by the multiplier replicates will be. The latter aspect will be discussed in detail in Section 4.5, where two broad classes of smooth estimators will be introduced and studied both theoretically and empirically.

#### 4.4.4 Smooth sequential dependent multiplier replicates

We now consider a similar construction but based on smooth analogs of  $\hat{\mathbb{B}}_n^{[b]}$  in (4.22) and  $\check{\mathbb{B}}_n^{[b]}$  in (4.23). Specifically, Corollary 4.2.13 suggests that, to bootstrap  $\mathbb{C}_n^\nu$  in (4.10), a first step is to bootstrap  $\hat{\mathbb{B}}_n^\nu$  in (4.15) or  $\check{\mathbb{B}}_n^\nu$  in (4.16). By analogy with (4.3) and (4.6), natural smooth analogs of  $\hat{\mathbb{B}}_n^{[b]}$  and  $\check{\mathbb{B}}_n^{[b]}$  could be defined, for any  $b \in \mathbb{N}$ ,  $(s, t) \in \Lambda$  and  $\mathbf{u} \in [0, 1]^d$ , by

$$\begin{aligned} \hat{\mathbb{B}}_n^{[b],\nu}(s, t, \mathbf{u}) &= \int_{[0,1]^d} \hat{\mathbb{B}}_n^{[b]}(s, t, \mathbf{w}) d\nu_{\mathbf{u}}^{\mathbf{X}_{1:n}}(\mathbf{w}) \\ &= \frac{1}{\sqrt{n}} \sum_{i=[ns]+1}^{[nt]} \xi_{i,n}^{[b]} \left\{ \mathcal{K}_{R_i^{1:n}}^{\mathbf{X}_{1:n}}(\mathbf{u}) - C_{1:n}^\nu(\mathbf{u}) \right\} \end{aligned} \quad (4.26)$$

and

$$\begin{aligned}\check{\mathbb{B}}_n^{[b],\nu}(s, t, \mathbf{u}) &= \int_{[0,1]^d} \check{\mathbb{B}}_n^{[b]}(s, t, \mathbf{w}) d\nu_{\mathbf{u}}^{\mathcal{X}_{[ns]+1:[nt]}}(\mathbf{w}) \\ &= \frac{1}{\sqrt{n}} \sum_{i=[ns]+1}^{[nt]} \xi_{i,n}^{[b]} \left\{ \mathcal{K}_{\mathbf{R}_i^{\mathcal{X}_{[ns]+1:[nt]}}}^{\mathcal{X}_{[ns]+1:[nt]}}(\mathbf{u}) - C_{[ns]+1:[nt]}^{\nu}(\mathbf{u}) \right\},\end{aligned}\quad (4.27)$$

respectively, where  $\mathcal{K}_{\mathbf{R}_i^{\mathcal{X}_{1:n}}}^{\mathcal{X}_{1:n}}$  and  $\mathcal{K}_{\mathbf{R}_i^{\mathcal{X}_{[ns]+1:[nt]}}}^{\mathcal{X}_{[ns]+1:[nt]}}$  are defined in (4.5). Combining these ingredients with estimators of the unknown partial derivatives of  $C$ , as smooth analogs of  $\hat{\mathbb{C}}_n^{[b]}$  in (4.24) and  $\check{\mathbb{C}}_n^{[b]}$  in (4.25), we obtain

$$\hat{\mathbb{C}}_n^{[b],\nu}(s, t, \mathbf{u}) = \hat{\mathbb{B}}_n^{[b],\nu}(s, t, \mathbf{u}) - \sum_{j=1}^d \dot{C}_{j,1:n}(\mathbf{u}) \hat{\mathbb{B}}_n^{[b],\nu}(s, t, \mathbf{u}^{(j)}), \quad (4.28)$$

and

$$\check{\mathbb{C}}_n^{[b],\nu}(s, t, \mathbf{u}) = \check{\mathbb{B}}_n^{[b],\nu}(s, t, \mathbf{u}) - \sum_{j=1}^d \dot{C}_{j,[ns]+1:[nt]}(\mathbf{u}) \check{\mathbb{B}}_n^{[b],\nu}(s, t, \mathbf{u}^{(j)}), \quad (4.29)$$

respectively, for  $b \in \mathbb{N}$ ,  $(s, t) \in \Lambda$  and  $\mathbf{u} \in [0, 1]^d$ .

To establish the asymptotic validity of these smooth multiplier bootstrap replicates, it will suffice that the partial derivative estimators satisfy the following rather natural mild condition.

**Condition 4.4.1** (Bounded and weakly consistent partial derivative estimators). There exists a constant  $\zeta > 0$  such that, for any  $j \in \{1, \dots, d\}$  and  $n \in \mathbb{N}$ ,

$$\sup_{(s,t,\mathbf{u}) \in \Lambda \times [0,1]^d} \left| \dot{C}_{j,[ns]+1:[nt]}(\mathbf{u}) \right| \leq \zeta.$$

Furthermore, for any  $\delta \in (0, 1)$ ,  $\varepsilon \in (0, 1/2)$  and  $j \in \{1, \dots, d\}$ ,

$$\sup_{\substack{(s,t) \in \Lambda \\ t-s \geq \delta}} \sup_{\substack{\mathbf{u} \in [0,1]^d \\ u_j \in [\varepsilon, 1-\varepsilon]}} \left| \dot{C}_{j,[ns]+1:[nt]}(\mathbf{u}) - \dot{C}_j(\mathbf{u}) \right| = o_{\mathbb{P}}(1).$$

In addition, following Bücher et al. (2014), we impose the following condition on the observations and the underlying multiplier sequences.

**Condition 4.4.2** (Strong mixing and multiplier conditions). One of the following two conditions holds:

- (i) The random vectors in  $\mathcal{X}_{1:n}$  are i.i.d. and  $(\xi_{i,n}^{[1]})_{i \in \mathbb{Z}}, (\xi_{i,n}^{[2]})_{i \in \mathbb{Z}}, \dots$  are independent copies of a multiplier sequence satisfying (M0).
- (ii) The stretch  $\mathcal{X}_{1:n}$  is drawn from a stationary sequence  $(\mathbf{X}_i)_{i \in \mathbb{Z}}$  whose strong mixing coefficients satisfy  $\alpha_r^{\mathbf{X}} = O(r^{-a})$  for some  $a > 3 + 3d/2$  as  $r \rightarrow \infty$ . Furthermore,

$(\xi_{i,n}^{[1]})_{i \in \mathbb{Z}}, (\xi_{i,n}^{[2]})_{i \in \mathbb{Z}}, \dots$  are independent copies of a dependent multiplier sequence satisfying (M1)–(M3) with  $\ell_n = O(n^{1/2-\gamma})$  for some  $0 < \gamma < 1/2$ .

The following result is proven in Appendix 4.7.4.

**Theorem 4.4.3** (Asymptotic validity of the smooth dependent multiplier bootstraps). *Under Conditions 4.2.8, 4.2.9, 4.4.1 and 4.4.2, for any  $b \in \mathbb{N}$ , there holds*

$$\sup_{(s,t,\mathbf{u}) \in \Lambda \times [0,1]^d} \left| \hat{\mathbb{C}}_n^{[b],\nu}(s,t,\mathbf{u}) - \hat{\mathbb{C}}_n^{[b]}(s,t,\mathbf{u}) \right| = o_{\mathbb{P}}(1), \quad (4.30)$$

$$\sup_{(s,t,\mathbf{u}) \in \Lambda \times [0,1]^d} \left| \check{\mathbb{C}}_n^{[b],\nu}(s,t,\mathbf{u}) - \check{\mathbb{C}}_n^{[b]}(s,t,\mathbf{u}) \right| = o_{\mathbb{P}}(1). \quad (4.31)$$

Furthermore,

$$\begin{aligned} (\mathbb{C}_n^\nu, \hat{\mathbb{C}}_n^{[1],\nu}, \hat{\mathbb{C}}_n^{[2],\nu}) &\rightsquigarrow (\mathbb{C}_C, \mathbb{C}_C^{[1]}, \mathbb{C}_C^{[2]}), \\ (\mathbb{C}_n^\nu, \check{\mathbb{C}}_n^{[1],\nu}, \check{\mathbb{C}}_n^{[2],\nu}) &\rightsquigarrow (\mathbb{C}_C, \mathbb{C}_C^{[1]}, \mathbb{C}_C^{[2]}) \end{aligned}$$

in  $\{\ell^\infty(\Lambda \times [0,1]^d)\}^3$ , where  $\mathbb{C}_C^{[1]}$  and  $\mathbb{C}_C^{[2]}$  are independent copies of  $\mathbb{C}_C$  defined in (4.14).

#### 4.4.5 Finite-sample comparison of three multiplier bootstraps

From Theorem 4.2.10, we know that, under Conditions 4.2.8 and 4.2.9, the classical sequential empirical copula process  $\mathbb{C}_n$  in (4.9) and the smooth sequential empirical copula process  $\mathbb{C}_n^\nu$  in (4.10) are asymptotically equivalent. In a related way, Theorem 4.4.3 provides conditions under which corresponding multiplier and smooth multiplier replicates are asymptotically equivalent. Although one expects that  $\mathbb{C}_n^\nu$  is probably best resampled using multiplier replicates constructed with the same smoothing distributions, that is, with  $\hat{\mathbb{C}}_n^{[b],\nu}$  in (4.28) or  $\check{\mathbb{C}}_n^{[b],\nu}$  in (4.29), we have no asymptotic results to support this (see also Remark 4.2.14). Indeed, given that all versions of multiplier replicates are asymptotically equivalent, it may well be that, for instance, in some cases, classical (non-smooth) multiplier replicates are equivalent or even preferable to smooth multiplier replicates when it comes to resampling  $\mathbb{C}_n^\nu$ . It is the aim of this section to study this empirically. For simplicity, we restrict our investigations to a non-sequential setting and independent observations.

Specifically, we designed experiments to study which multiplier replicates are best suited to estimate certain functionals of the three (non-sequential) empirical copula processes defined, for any  $\mathbf{u} \in [0,1]^d$ , by

$$\mathbb{C}_n^{\text{Dirac}}(\mathbf{u}) = \sqrt{n}\{C_{1:n}(\mathbf{u}) - C(\mathbf{u})\} = \mathbb{C}_n(0,1,\mathbf{u}), \quad (4.32)$$

$$\mathbb{C}_n^{\text{Bin}}(\mathbf{u}) = \sqrt{n}\{C_{1:n}^{\text{Bin}}(\mathbf{u}) - C(\mathbf{u})\}, \quad (4.33)$$

$$\mathbb{C}_n^{\text{BetaB4}}(\mathbf{u}) = \sqrt{n}\{C_{1:n}^{\text{BetaB4}}(\mathbf{u}) - C(\mathbf{u})\}, \quad (4.34)$$

where

- $\mathbb{C}_n$  is the classical (non-smooth) sequential empirical copula process defined in (4.9),
- $C_{1:n}^{\text{Bin}}$  is the empirical beta copula in (4.17) (which is obtained by considering smoothing distributions with scaled binomial margins and independence copula as explained in Section 4.2.1),
- $C_{1:n}^{\text{BetaB4}}$  is the version of  $C_{1:n}^\nu$  introduced in Section 4.2.1 obtained by considering smoothing distributions with scaled beta-binomial margins and survival copula the empirical beta copula  $C_{1:n}^{\text{Bin}}$ , and found to have the best finite-sample performance in the numerical experiments of Kojadinovic and Yi (2022).

As already mentioned, since we are in a non-sequential setting, the two generic multiplier replicates defined in (4.28) and (4.29) coincide. To approximate the “sampling distributions” of the three empirical copula processes defined above, we considered as candidate bootstraps the multiplier replicates defined using the same smoothing distributions. They will be denoted by  $\hat{\mathbb{C}}_n^{[b],\text{Dirac}}$ ,  $\hat{\mathbb{C}}_n^{[b],\text{Bin}}$  and  $\hat{\mathbb{C}}_n^{[b],\text{BetaB4}}$ ,  $b \in \mathbb{N}$ , respectively, as we continue. To only investigate the effect of the choice of the smoothing distributions involved in the definition of  $\hat{\mathbb{B}}_n^{[b],\nu}$  in (4.26), all three multiplier replicates were computed using the true partial derivative  $\dot{C}_j$ ,  $j \in \{1, \dots, d\}$ . Furthermore, since we restricted our experiments to independent observations, all the multiplier replicates were based on i.i.d. multiplier sequences defined in (M0) in Section 4.4.2. Following Bücher and Dette (2010), these sequences were simply taken to be random samples drawn from the uniform distribution on  $\{-1, 1\}$ .

For the design of our experiments, we followed again Bücher and Dette (2010). First, for  $d = 2$ , we assessed how well the covariances of the empirical processes  $\mathbb{C}_n^{\text{Dirac}}$  in (4.32),  $\mathbb{C}_n^{\text{Bin}}$  in (4.33) and  $\mathbb{C}_n^{\text{BetaB4}}$  in (4.34) at the points  $P = \{(i/3, j/3) : i, j = 1, 2\}$  can be approximated using the three possible multiplier bootstrap replicates. For each target empirical copula process, we began by precisely estimating its covariance at the points in  $P$  from 100 000 independent samples of size  $n \in \{10, 20, 40, 80\}$  drawn from a bivariate copula  $C$  with a Kendall’s tau of  $\tau \in \{0, 0.25, 0.5, 0.75\}$ . For  $C$ , we considered either the Clayton or the Gumbel–Hougaard copula. Next, for each considered combination of  $C$ ,  $n$ ,  $\tau$ , target process and multiplier process, we generated 1000 samples from  $C$ , and, for each sample, we computed  $B = 1000$  multiplier bootstrap replicates. These  $B = 1000$  replicates were used to obtain one estimate of the covariance of the target process at the points in  $P$ .

The results when  $C$  is the Clayton copula with a Kendall’s tau of  $\tau \in \{0, 0.25, 0.5, 0.75\}$  are reported in Figure 4.1. The first (resp. second, third) column of graphs reports the average of the empirical mean square errors (MSEs)  $\times 10^4$  of the three candidate multiplier estimators of the covariance of  $\mathbb{C}_n^{\text{Dirac}}$  (resp.  $\mathbb{C}_n^{\text{Bin}}$ ,  $\mathbb{C}_n^{\text{BetaB4}}$ ) at the points in  $P$  against the sample size  $n$ . Each row of graphs corresponds to a different value of  $\tau$ . In the top-left panel for instance, the solid (resp. dashed, dotted) curve gives the average MSE when the covariance of  $\mathbb{C}_n^{\text{Dirac}}$  is estimated using  $\hat{\mathbb{C}}_n^{[b],\text{Dirac}}$  (resp.  $\hat{\mathbb{C}}_n^{[b],\text{Bin}}$ ,  $\hat{\mathbb{C}}_n^{[b],\text{BetaB4}}$ ).



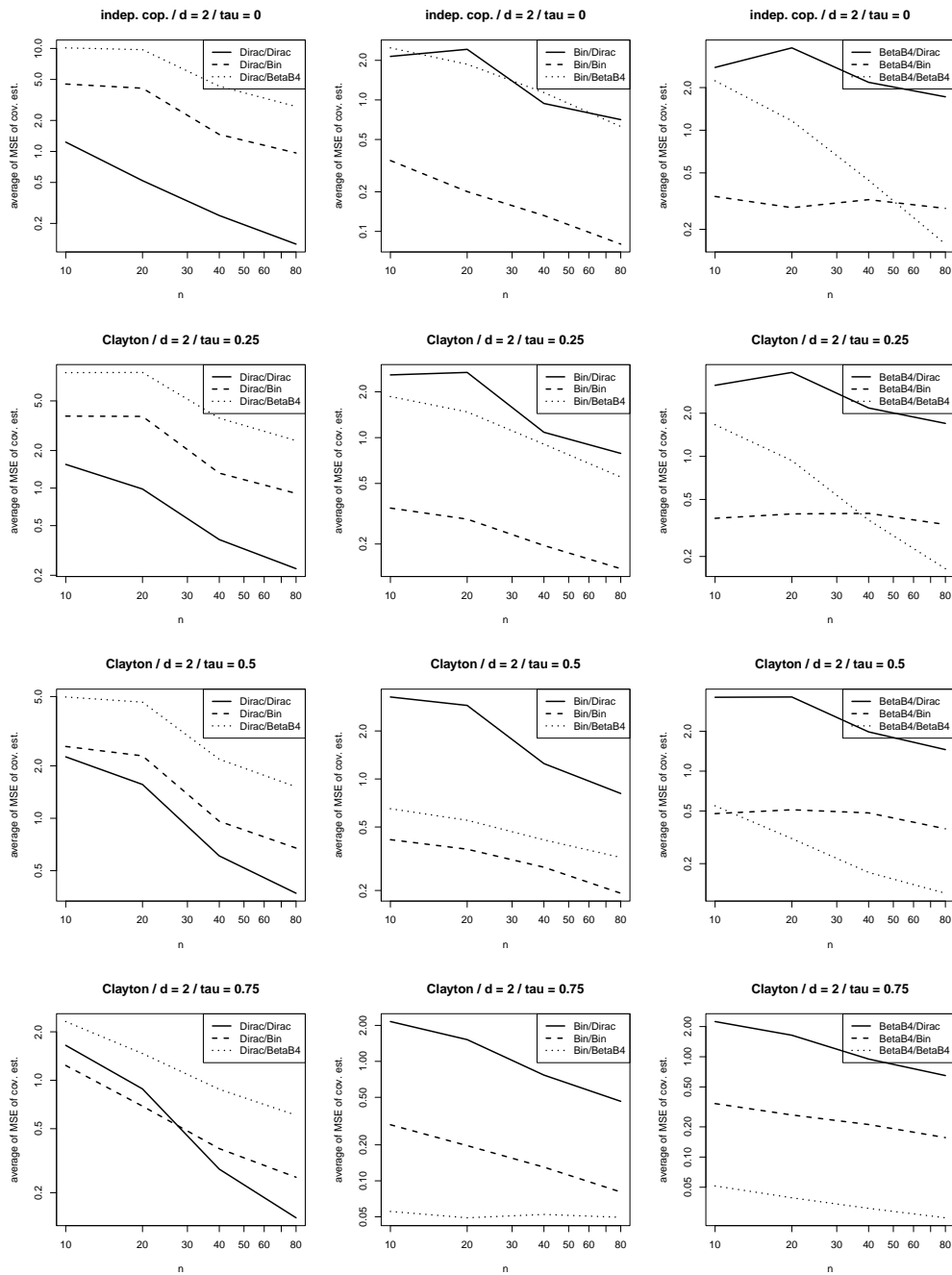


FIGURE 4.1: For observations generated from the bivariate Clayton copula with a Kendall's tau of  $\tau \in \{0, 0.25, 0.5, 0.75\}$  and for each combination of target and multiplier process, average of the empirical MSEs ( $\times 10^4$ ) of the bootstrap estimators of the covariance of the target process at the points in  $P$  against the sample size  $n$ . The legend "Dirac/Bin" for instance refers to the situation when the target process is  $C_n^{\text{Dirac}}$  and the multiplier process is  $\hat{C}_n^{[b], \text{Bin}}$ .

As one can see, reassuringly, all the curves are globally decreasing, confirming that, for each target process, the bootstrap approximations improve as  $n$  increases. A more careful inspection reveals that, in almost all settings, it is the multiplier bootstrap constructed with the same smoothing distributions as the target process that leads to the best estimation. It is actually

only when  $\mathbb{C}_n^{\text{BetaB4}}$  is the target process that covariance estimations based on  $\hat{\mathbb{C}}_n^{[b],\text{Bin}}$  are sometimes better than estimations based on  $\hat{\mathbb{C}}_n^{[b],\text{BetaB4}}$ . This happens mostly for small  $n$  and  $\tau$ . Results for the Gumbel–Hougaard copula (not reported) are not qualitatively different.

In a second experiment, we assessed how well high quantiles of

$$KS(f_n) = \sup_{\mathbf{u} \in [0,1]^d} |f_n(\mathbf{u})| \quad \text{and} \quad CvM(f_n) = \int_{[0,1]^d} \{f_n(\mathbf{u})\}^2 d\mathbf{u} \quad (4.35)$$

for  $d \in \{2, 3\}$  and  $f_n \in \{\mathbb{C}_n^{\text{Dirac}}, \mathbb{C}_n^{\text{Bin}}, \mathbb{C}_n^{\text{BetaB4}}\}$  can be estimated by the three candidate multiplier bootstraps. From a practical perspective, the integral in (4.35) was approximated by a mean using a uniform grid on  $(0, 1)^d$  of size  $10^2$  when  $d = 2$  and  $5^3$  when  $d = 3$ . For  $d \in \{2, 3\}$ ,  $C$  the Clayton or the Gumbel-Hougaard copula whose bivariate margins have a Kendall’s tau of  $\tau \in \{0, 0.25, 0.5, 0.75\}$  and  $n \in \{10, 20, 40, 80\}$ , the 90% and 95%-quantiles of  $CvM(f_n)$  were first precisely estimated from 100 000 independent samples of size  $n$  drawn from  $C$ . Next, for each combination of  $d$ ,  $C$ ,  $n$ ,  $\tau$ , target process and multiplier process, we generated 1000 samples from  $C$  and, for each sample, we computed  $B = 1000$  multiplier bootstrap replicates. These  $B = 1000$  replicates were used to obtain one estimate of each of the target quantiles. Following [Kojadinovic and Stemikovskaya \(2019\)](#), all such estimations were carried out using *centered* replicates of  $f_n$ . When  $f_n = \mathbb{C}_n^{\text{Dirac}}$  for instance, this amounts to using, for any  $\mathbf{u} \in [0, 1]^d$  and  $b \in \{1, \dots, B\}$ ,

$$\hat{\mathbb{C}}_n^{[b],\text{Dirac}}(\mathbf{u}) - \frac{1}{B} \sum_{b=1}^B \hat{\mathbb{C}}_n^{[b],\text{Dirac}}(\mathbf{u}),$$

instead of  $\hat{\mathbb{C}}_n^{[b],\text{Dirac}}(\mathbf{u}) = \hat{\mathbb{C}}_n^{[b]}(0, 1, \mathbf{u})$  in (4.24). The centered versions of the other replicates are defined analogously. The rationale behind centering is that the replicates, whatever their type, can be regarded as computable approximations of the limiting *centered* Gaussian process  $\mathbb{C}_C(0, 1, \cdot)$  in (4.14); see, for instance, Theorem 4.4.3. Note that the use of centered replicates was found to always lead to better finite-sample performance in the related Monte Carlo experiments carried out in [Kojadinovic and Stemikovskaya \(2019\)](#). Its use is however irrelevant in the previous covariance estimation experiment given the formula of the empirical covariance.

The results for the 95%-quantiles of the Kolmogorov–Smirnov functionals when  $C$  is the trivariate Gumbel–Hougaard are reported in Figure 4.2. The conclusions are overall similar to those obtained after the first experiment:

- The 95%-quantile of the Kolmogorov–Smirnov functional of  $\mathbb{C}_n^{\text{Dirac}}$  is always best estimated using the corresponding empirical quantile of the same functional of  $\hat{\mathbb{C}}_n^{[b],\text{Dirac}}$ .
- When the target process is  $\mathbb{C}_n^{\text{Bin}}$ , the best results are obtained when the multiplier process is  $\hat{\mathbb{C}}_n^{[b],\text{Bin}}$ , except in the case of strongly dependent observations in which case, for the sample sizes under consideration,  $\hat{\mathbb{C}}_n^{[b],\text{Dirac}}$  gives better estimations.

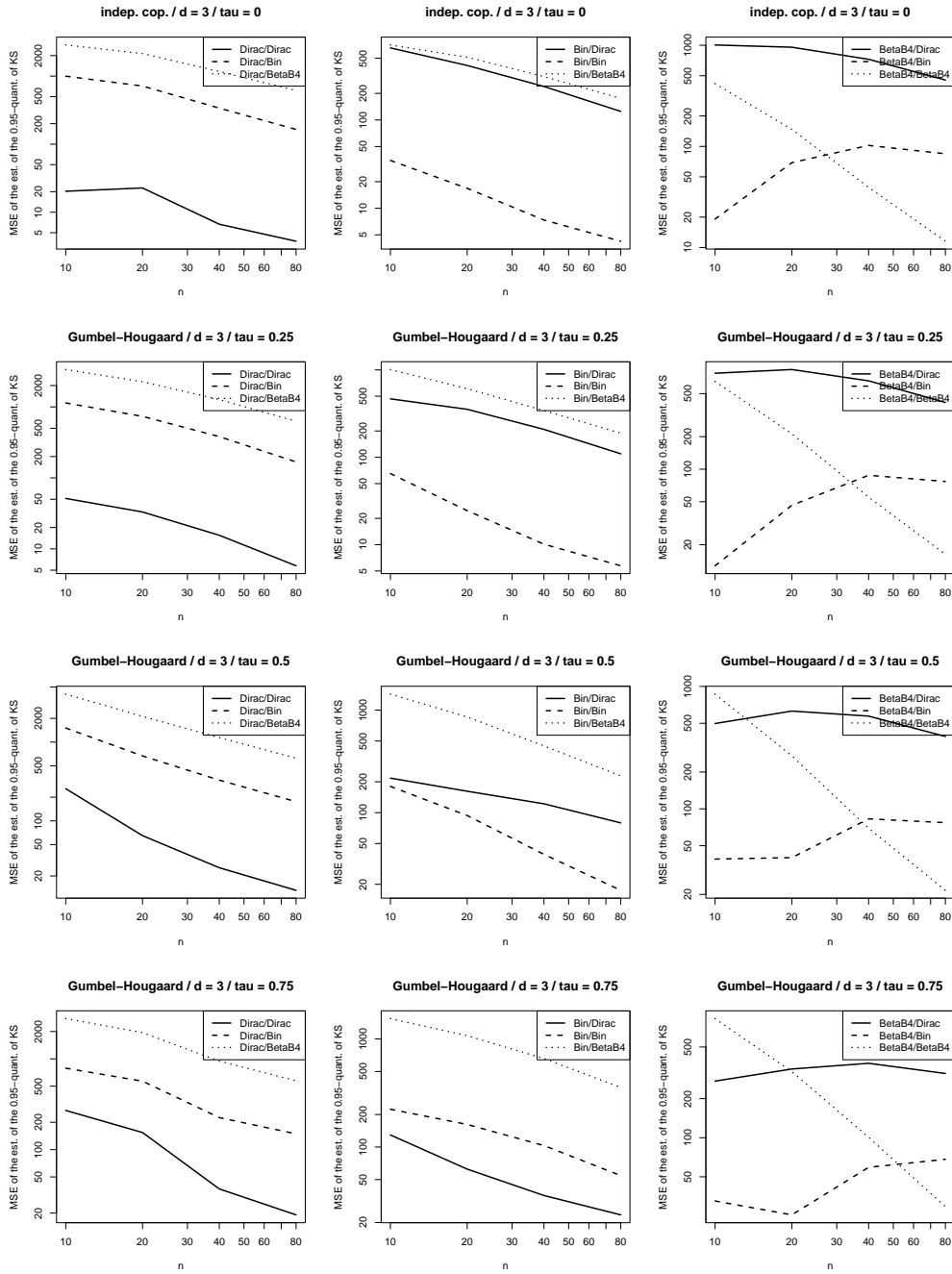


FIGURE 4.2: For observations generated from the trivariate Gumbel–Hougaard copula whose bivariate margins have a Kendall’s tau of  $\tau \in \{0, 0.25, 0.5, 0.75\}$ , empirical MSE ( $\times 10^4$ ) of the three candidate multiplier estimators of high quantiles of  $KS(f_n)$  in (4.35) for  $f_n \in \{\mathbb{C}_n^{\text{Dirac}}, \mathbb{C}_n^{\text{Bin}}, \mathbb{C}_n^{\text{BetaB4}}\}$  against the sample size  $n$ . The legend “Dirac/Bin” for instance refers to the situation when the target process is  $\mathbb{C}_n^{\text{Dirac}}$  and the multiplier process is  $\hat{\mathbb{C}}_n^{[b],\text{Bin}}$ .

- When the target process is  $\mathbb{C}_n^{\text{BetaB4}}$ , it is only when  $n$  reaches 40 or 80 that the best estimations are obtained using  $\hat{\mathbb{C}}_n^{[b],\text{BetaB4}}$ . For smaller  $n$ , the use of  $\hat{\mathbb{C}}_n^{[b],\text{Bin}}$  gives better results.

Results for the Clayton copula, 90%-quantiles, dimension  $d = 2$  or Cramér–von Mises functionals (not reported) are not qualitatively different.

The previous experiments confirm that it seems meaningful to resample  $\mathbb{C}_n^\nu$  in (4.10) using multiplier replicates constructed with the same smoothing distributions, that is, with  $\hat{\mathbb{C}}_n^{[b],\nu}$  in (4.28) or  $\check{\mathbb{C}}_n^{[b],\nu}$  in (4.29), although this choice may not be optimal in certain cases when  $n$  is small.

#### 4.4.6 Application to change-point detection

A natural application area for the smooth sequential empirical copula process  $\mathbb{C}_n^\nu$  in (4.10) is that of change-point detection. To illustrate the possible advantages coming from the use of smooth empirical copulas in inference procedures, we first briefly explain in this section how the previous derivations can be used to obtain a smooth version of the test proposed in Bücher et al. (2014) for detecting changes in the cross-sectional dependence of multivariate time series. We then reproduce some of the experiments of Bücher et al. (2014) to compare the (non-smooth) test proposed therein with its smooth version based on the empirical beta copula and on corresponding smooth bootstrap replicates. Note that we did not consider the use of the alternative data-adaptive smoothing distributions considered in Kojadinovic and Yi (2022) and leading to the estimator  $C_{k:l}^{\text{BetaB4}}$  because they incur a substantially higher computational cost.

The null hypothesis of such tests is that  $\mathbf{X}_{1:n}$  is a stretch from a stationary time series (of continuous random vectors) and their aim is to be particularly sensitive to the alternative hypothesis

$$H_1 : \exists \text{ distinct } C_1, C_2 \text{ and } k^* \in \{1, \dots, n-1\} \text{ such that} \tag{4.36}$$

$$\mathbf{X}_1, \dots, \mathbf{X}_{k^*} \text{ have copula } C_1 \text{ and } \mathbf{X}_{k^*+1}, \dots, \mathbf{X}_n \text{ have copula } C_2.$$

The ingredients of the smooth version of the test can be obtained *mutatis mutandis* from Bücher et al. (2014). Specifically, we consider as test statistic the maximally selected Cramér–von Mises functional defined by

$$S_n^\nu = \sup_{s \in [0,1]} \int_{[0,1]^d} \{\mathbb{D}_n^\nu(s, \mathbf{u})\}^2 dC_{1:n}(\mathbf{u}),$$

where

$$\mathbb{D}_n^\nu(s, \mathbf{u}) = \sqrt{n} \lambda_n(0, s) \lambda_n(s, 1) \{C_{1:[ns]}^\nu(\mathbf{u}) - C_{[ns]+1:n}^\nu(\mathbf{u})\}, \quad (s, \mathbf{u}) \in [0, 1]^{d+1}.$$

As one can see, the latter involves comparisons of (smooth) empirical copulas computed from subsamples of the data. Noticing that, under the null,

$$\mathbb{D}_n^\nu(s, \mathbf{u}) = \lambda_n(s, 1) \mathbb{C}_n^\nu(0, s, \mathbf{u}) - \lambda_n(0, s) \mathbb{C}_n^\nu(s, 1, \mathbf{u}), \quad (s, \mathbf{u}) \in [0, 1]^{d+1},$$

possible multiplier bootstrap replicates for  $S_n^\nu$  can be defined either by

$$\hat{S}_n^{[b],\nu} = \sup_{s \in [0,1]} \int_{[0,1]^d} \{\hat{\mathbb{D}}_n^{[b],\nu}(s, \mathbf{u})\}^2 dC_{1:n}(\mathbf{u}), \quad b \in \mathbb{N},$$

or by

$$\check{S}_n^{[b],\nu} = \sup_{s \in [0,1]} \int_{[0,1]^d} \{\check{\mathbb{D}}_n^{[b],\nu}(s, \mathbf{u})\}^2 dC_{1:n}(\mathbf{u}), \quad b \in \mathbb{N}, \quad (4.37)$$

where, for any  $(s, \mathbf{u}) \in [0, 1]^{d+1}$ ,

$$\begin{aligned} \hat{\mathbb{D}}_n^{[b],\nu}(s, \mathbf{u}) &= \lambda_n(s, 1) \hat{\mathbb{C}}_n^{[b],\nu}(0, s, \mathbf{u}) - \lambda_n(0, s) \hat{\mathbb{C}}_n^{[b],\nu}(s, 1, \mathbf{u}), \\ \check{\mathbb{D}}_n^{[b],\nu}(s, \mathbf{u}) &= \lambda_n(s, 1) \check{\mathbb{C}}_n^{[b],\nu}(0, s, \mathbf{u}) - \lambda_n(0, s) \check{\mathbb{C}}_n^{[b],\nu}(s, 1, \mathbf{u}), \end{aligned}$$

with  $\hat{\mathbb{C}}_n^{[b],\nu}$  and  $\check{\mathbb{C}}_n^{[b],\nu}$  defined in (4.28) and (4.29), respectively. Note that, in the expressions of the multiplier replicates of  $C_n^\nu$ , as estimators of the first-order partial derivatives of the copula, we use the ‘‘truncated’’ finite-difference based estimators defined in (4.45) of the forthcoming section with bandwidths  $h = h' = \min\{(l - k + 1)^{-1/2}, 1/2\}$ . As we will see from Proposition 4.5.5, the latter can satisfy Condition 4.4.1. Finally, as in Bücher et al. (2014), approximate p-values for  $S_n^\nu$  can be computed via either

$$\frac{1}{B} \sum_{b=1}^B \mathbf{1}(\hat{S}_n^{[b],\nu} \geq S_n^\nu) \quad \text{or} \quad \frac{1}{B} \sum_{b=1}^B \mathbf{1}(\check{S}_n^{[b],\nu} \geq S_n^\nu),$$

for some large integer  $B$ . Theoretical results confirming that the above way of proceeding is asymptotically valid under the null can be obtained by starting from Theorem 4.4.3, proceeding as in Bücher et al. (2014) and finally using results stated in Section 4 of Bücher and Kojadinovic (2019).

If, for any  $m \in \mathbb{N}$ , the underlying smoothing distributions  $\nu_{\mathbf{x}}^{\mathbf{u}}$ ,  $\mathbf{x} \in (\mathbb{R}^d)^m$ ,  $\mathbf{u} \in [0, 1]^d$ , are Dirac measures at  $\mathbf{u}$ , the previous ingredients are non-smooth and the resulting test coincides exactly with the test studied in Bücher et al. (2014). The test statistic will naturally be denoted by  $S_n^{\text{Dirac}}$  in that case. As alternative smoothing distributions, we considered those leading to the empirical beta copula and specified in Remark 4.2.2 as well as at the end of Section 4.2.1. The resulting statistic will then naturally be denoted by  $S_n^{\text{Bin}}$ .

To compare the test based on  $S_n^{\text{Bin}}$  to the test based on  $S_n^{\text{Dirac}}$ , we considered experiments similar to those reported in Section 5 of Bücher et al. (2014). Both tests were carried out at the 5% significance level using replicates of the form (4.37) as these seemed to lead to better results. The dependent multiplier sequences necessary to carry out the tests were generated as explained in the last paragraph of Appendix C of Bücher et al. (2014). The value of the bandwidth parameter  $\ell_n$  appearing in (M2) and (M3) in Section 4.4.2 was chosen using the procedure described in Bücher and Kojadinovic (2016, Section 5) (although this way of proceeding may not be ‘‘optimal’’ for the smooth multiplier bootstrap replicates).

TABLE 4.3: Percentages of rejection of the null hypothesis of stationarity computed from 1000 samples of size  $n \in \{25, 50, 100, 200\}$  generated as explained in Section 4.4.6, where  $C$  is the bivariate Frank copula with a Kendall's tau of  $\tau \in \{0, 0.33, 0.66\}$ .

$\beta$	$\tau$	$n = 25$		$n = 50$		$n = 100$		$n = 200$	
		$S_n^{\text{Dirac}}$	$S_n^{\text{Bin}}$	$S_n^{\text{Dirac}}$	$S_n^{\text{Bin}}$	$S_n^{\text{Dirac}}$	$S_n^{\text{Bin}}$	$S_n^{\text{Dirac}}$	$S_n^{\text{Bin}}$
0	0.00	17.5	13.3	7.7	8.0	5.5	5.8	3.8	4.4
	0.33	18.7	13.4	7.6	7.3	4.9	6.3	4.2	4.0
	0.66	21.1	11.7	5.6	4.9	3.0	3.1	3.2	3.8
0.3	0.00	18.8	16.1	6.2	7.4	4.3	4.7	6.4	6.0
	0.33	21.4	16.4	7.8	8.7	5.2	5.9	5.4	5.4
	0.66	25.3	16.8	5.4	5.9	2.1	3.0	1.2	1.4
0.5	0.00	26.1	22.8	11.4	11.7	6.1	6.6	6.2	7.2
	0.33	22.9	23.0	10.3	11.2	5.5	7.2	2.4	3.6
	0.66	27.5	20.1	10.5	11.0	2.2	3.6	1.6	1.6

As a first experiment, we estimated the percentages of rejection of the null hypothesis of stationarity for data generated under the null. As data generating model, we used a bivariate AR(1) model. Specifically, let  $\mathbf{U}_i$ ,  $i \in \{-100, \dots, n\}$ , be a bivariate i.i.d. sample from a copula  $C$ . Then, set  $\epsilon_i = (\Phi^{-1}(U_{i1}), \Phi^{-1}(U_{i2}))$ , where  $\Phi$  is the d.f. of the standard normal distribution, and  $\mathbf{X}_{-100} = \epsilon_{-100}$ . Finally, for any  $j \in \{1, 2\}$  and  $i \in \{-99, \dots, n\}$ , compute recursively

$$X_{ij} = \beta X_{i-1,j} + \epsilon_{ij},$$

where the first 100 observations are used as a burn-out sample.

We considered  $n \in \{25, 50, 100, 200\}$ ,  $C$  to be bivariate Frank copula with a Kendall's tau of  $\tau \in \{0, 0.33, 0.66\}$  and  $\beta \in \{0, 0.3, 0.5\}$ . The corresponding rejection percentages are reported in Table 4.3. As one can see, both tests appear to hold their level reasonably well when  $n \in \{100, 200\}$ . The tests should however clearly not be used when  $n = 25$  but might be employed when  $n = 50$  in the case of weakly serially dependent data.

As a second experiment, we estimated rejection percentages of the null hypothesis of stationarity for data generated under  $H_1$  in (4.36). To do so, we considered a similar data generating model as in the first experiment except that the  $\mathbf{U}_i$ 's for  $i \in \{-100, \dots, k^*\}$  are i.i.d. from a copula  $C_1$  while the  $\mathbf{U}_i$ 's for  $i \in \{k^* + 1, \dots, n\}$  are i.i.d. from a copula  $C_2 \neq C_1$ . Following Bücher et al. (2014), we set  $k^* = \lfloor nt \rfloor$  with  $t \in \{0.1, 0.25, 0.5\}$  and considered  $n \in \{50, 100, 200\}$ ,  $C_1$  the bivariate Frank copula with a Kendall's tau 0.2 and  $C_2$  the bivariate Frank copula with a Kendall's tau in  $\{0.4, 0.6\}$ . The results are reported in Table 4.4. As one can see, the test based on  $S_n^{\text{Bin}}$  appears overall to be more powerful than the one based on  $S_n^{\text{Dirac}}$ . The largest differences in power tend to occur for  $\tau = 0.6$  and  $t \in \{0.1, 0.25\}$  which corresponds to the situation when the test statistic should be the largest because of a difference between an empirical copula computed from a small number of observations (approximately  $\lfloor nt \rfloor$ ) and an empirical copula computed from the remaining observations. While one cannot conclude that smooth change-point detection tests such as the one based on  $S_n^{\text{Bin}}$  will be more powerful than the non-smooth test based on  $S_n^{\text{Dirac}}$  in all situations, the obtained results confirm in part the intuition that

TABLE 4.4: Percentages of rejection of the null hypothesis of stationarity computed from 1000 samples of size  $n \in \{50, 100, 200\}$  generated under  $H_1$  as explained in Section 4.4.6, where  $k^* = \lfloor nt \rfloor$ ,  $C_1$  and  $C_2$  are both bivariate Frank copulas such that  $C_1$  has a Kendall's tau of 0.2 and  $C_2$  a Kendall's tau of  $\tau \in \{0.4, 0.6\}$ .

$\tau$	$n$	$t$	$\beta = 0$		$\beta = 0.3$	
			$S_n^{\text{Dirac}}$	$S_n^{\text{Bin}}$	$S_n^{\text{Dirac}}$	$S_n^{\text{Bin}}$
0.4	50	0.10	8.8	8.7	8.0	8.1
		0.25	13.5	16.1	14.0	15.5
		0.50	14.7	15.3	17.5	18.4
	100	0.10	4.0	4.9	5.5	7.6
		0.25	16.9	19.3	14.8	17.9
		0.50	26.6	28.8	22.5	25.3
	200	0.10	6.6	7.4	5.6	6.6
		0.25	29.4	31.8	22.0	24.2
		0.50	51.4	53.8	42.0	43.8
0.6	50	0.10	10.2	13.0	9.1	11.6
		0.25	33.0	39.8	31.6	39.1
		0.50	53.0	56.8	47.0	51.1
	100	0.10	12.1	16.6	8.6	12.5
		0.25	62.6	70.9	51.9	60.3
		0.50	83.1	84.9	75.0	78.6
	200	0.10	30.4	37.8	21.0	28.8
		0.25	95.2	97.0	87.8	91.0
		0.50	99.4	99.4	97.0	97.2

smooth tests might be more sensitive to changes at the beginning or at the end of the data sequence.

## 4.5 Estimators of the first-order partial derivatives of the copula

The multiplier bootstrap replicates defined in the previous section all depend on the choice of estimators of the first-order partial derivatives of  $C$ . For asymptotic reasons, the latter were required to satisfy Condition 4.4.1. Obviously, the more accurate such estimators, the better we can expect the multiplier bootstraps to behave, whether they involve smoothing or not. After recalling existing definitions of such estimators based on finite differences of the classical empirical copula, we define two related classes of smooth estimators. Then, upon an appropriate choice of the underlying bandwidth parameters, we establish their weak consistency in a sequential setting which implies that many of the considered estimators satisfy Condition 4.4.1. In the last subsection, we report the results of bivariate and trivariate Monte Carlo experiments comparing selected estimators in terms of integrated mean squared error.

Note that, as already mentioned in the introduction, the results of this section can be of independent interest since, as discussed for instance in Janssen et al. (2016), estimators of the first-order partial derivatives of a copula have applications in mean and quantile regression as they lead to estimators of the conditional distribution function. In particular, as we shall see,

several estimators considered in our Monte Carlo experiments display a better finite-sample performance than the Bernstein estimator studied in [Janssen et al. \(2016\)](#).

### 4.5.1 Estimators based on finite differences of the empirical copula

As already mentioned in Section 4.4.3, in their seminal work on the multiplier bootstrap for the classical empirical copula process, [Rémillard and Scaillet \(2009\)](#) considered estimators of the first-order partial derivatives  $\dot{C}_j$ ,  $j \in \{1, \dots, d\}$ , of  $C$  based on finite-differences of the empirical copula. In a sequential context, given a stretch  $\mathcal{X}_{k:l} = (\mathbf{X}_k, \dots, \mathbf{X}_l)$ ,  $1 \leq k \leq l \leq n$ , of observations and two bandwidth parameters  $h$  and  $h'$  in  $[0, 1/2]$  such that  $h + h' > 0$ , a slightly more general definition of the aforementioned estimators is

$$\dot{C}_{j,k:l,h,h'}^\nabla(\mathbf{u}) = \frac{C_{k:l}\{(\mathbf{u} + h\mathbf{e}_j) \wedge \mathbf{1}\} - C_{k:l}\{(\mathbf{u} - h'\mathbf{e}_j) \vee \mathbf{0}\}}{h + h'}, \quad \mathbf{u} \in [0, 1]^d, \quad (4.38)$$

where  $\mathbf{e}_j$  is the  $j$ th vector of the canonical basis of  $\mathbb{R}^d$ ,  $\mathbf{0} = (0, \dots, 0)$ ,  $\mathbf{1} = (1, \dots, 1) \in \mathbb{R}^d$ ,  $\wedge$  (resp.  $\vee$ ) denotes the minimum (resp. maximum) componentwise operator and  $C_{k:l}$  is the classical empirical copula of  $\mathcal{X}_{k:l}$  defined in (4.2). The symbol  $\nabla$  indicates that the estimators are based on finite-differences of  $C_{k:l}$  with “right” (resp. “left”) bandwidth  $h$  (resp.  $h'$ ).

In order to reduce the bias of the previous estimator for evaluation points  $\mathbf{u} \in [0, 1]$  with  $u_j \in [0, h') \cup (1 - h, 1]$ , [Kojadinovic et al. \(2011\)](#) considered the following minor variation of (4.38):

$$\dot{C}_{j,k:l,h,h'}^\Delta(\mathbf{u}) = \frac{C_{k:l}\{(\mathbf{u} + h\mathbf{e}_j) \wedge \mathbf{1}\} - C_{k:l}\{(\mathbf{u} - h'\mathbf{e}_j) \vee \mathbf{0}\}}{(u_j + h) \wedge 1 - (u_j - h') \vee 0}, \quad \mathbf{u} \in [0, 1]^d. \quad (4.39)$$

Note the use of the symbol  $\Delta$  still referring to finite-differences but upside-down compared to  $\nabla$  to distinguish (4.39) from (4.38).

As is well known, in general,  $\dot{C}_j$  exists almost everywhere on  $[0, 1]^d$  and, for those  $\mathbf{u} \in [0, 1]^d$  for which it exists,  $0 \leq \dot{C}_j(\mathbf{u}) \leq 1$  (see e.g., [Nelsen, 2006](#), Theorem 2.2.7). A natural modification of the estimators  $\dot{C}_{j,k:l,h,h'}^\nabla$  in (4.38) and  $\dot{C}_{j,k:l,h,h'}^\Delta$  in (4.39) thus consists of ensuring that they take their values in  $[0, 1]$  by truncating them:

$$\underline{\dot{C}}_{j,k:l,h,h'}^\nabla = (\dot{C}_{j,k:l,h,h'}^\nabla \vee 0) \wedge 1, \quad (4.40)$$

$$\underline{\dot{C}}_{j,k:l,h,h'}^\Delta = (\dot{C}_{j,k:l,h,h'}^\Delta \vee 0) \wedge 1. \quad (4.41)$$

Notice that taking the maximum with 0 in the previous expressions is actually not necessary as the estimators in (4.38) and (4.39) cannot be negative since the empirical copula  $C_{k:l}$  is a multivariate d.f. We nonetheless keep (4.40) and (4.41) as they are to be consistent with certain forthcoming definitions. More generally, in the rest of this section, underlining will be used to denote estimators constrained to take their values in  $[0, 1]$ .



### 4.5.2 Two classes of smooth estimators

To obtain smooth estimators of the first-order partial derivatives of  $C$ , the proposals in (4.38) and (4.39) can be extended in two natural ways. The first approach consists of considering finite-differences of smooth estimators of  $C$ . Given a stretch  $\mathcal{X}_{k:l}$ ,  $1 \leq k \leq l \leq n$ , of observations and two bandwidth parameters  $h$  and  $h'$  in  $[0, 1/2]$  such that  $h + h' > 0$ , this leads to the estimators

$$\dot{C}_{j,k:l,h,h'}^{\nu,\nabla}(\mathbf{u}) = \frac{C_{k:l}^{\nu}\{(\mathbf{u} + h\mathbf{e}_j) \wedge \mathbf{1}\} - C_{k:l}^{\nu}\{(\mathbf{u} - h'\mathbf{e}_j) \vee \mathbf{0}\}}{h + h'}, \quad \mathbf{u} \in [0, 1]^d, \quad (4.42)$$

$$\dot{C}_{j,k:l,h,h'}^{\nu,\Delta}(\mathbf{u}) = \frac{C_{k:l}^{\nu}\{(\mathbf{u} + h\mathbf{e}_j) \wedge \mathbf{1}\} - C_{k:l}^{\nu}\{(\mathbf{u} - h'\mathbf{e}_j) \vee \mathbf{0}\}}{(u_j + h) \wedge 1 - (u_j - h') \vee 0}, \quad \mathbf{u} \in [0, 1]^d, \quad (4.43)$$

where  $C_{k:l}^{\nu}$  is the smooth empirical copula of  $\mathcal{X}_{k:l}$  defined in (4.3). Notice the order of the symbols  $\nu$  and  $\nabla$  (resp.  $\Delta$ ) indicating that the empirical copula is first smoothed before finite-differencing is applied. Clearly, (4.38) (resp. (4.39)) is a particular case of (4.42) (resp. (4.43)) when the smoothing distributions  $\nu_{\mathbf{u}}^{\mathcal{X}_{k:l}}$ ,  $\mathbf{u} \in [0, 1]^d$ , in (4.3) are chosen to be Dirac measures at  $\mathbf{u} \in [0, 1]^d$ .

*Remark 4.5.1.* Since  $C$  is a multivariate d.f. with standard uniform margins, we have that, for any  $j \in \{1, \dots, d\}$  and  $\mathbf{u} \in V_j$  (where  $V_j$  is defined in Condition 4.2.8),  $\dot{C}_j(\mathbf{u}^{(j)}) = \lim_{h \rightarrow 0} \{C(\mathbf{u}^{(j)} + h\mathbf{e}_j) - C(\mathbf{u}^{(j)})\}/h = 1$  (where the notation  $\mathbf{u}^{(j)}$  is defined above Theorem 4.2.12). Interestingly enough, the estimator  $\dot{C}_{j,k:l,h,h'}^{\nu,\Delta}$  in (4.43) can satisfy this boundary constraint, that is, we can have  $\dot{C}_{j,k:l,h,h'}^{\nu,\Delta}(\mathbf{u}^{(j)}) = 1$ . This will indeed happen if  $C_{k:l}^{\nu}$  is a genuine copula, which according to Proposition 4.2.7, can occur under Condition 4.2.3 for specific choices of the smoothing distributions in (4.3) such as those leading to the empirical copulas  $C_{k:l}^{\text{Bin}}$  or  $C_{k:l}^{\text{BetaB4}}$  defined in Section 4.2.1.

By analogy with (4.40) and (4.41), it is straightforward to define truncated versions of the estimators in (4.42) and (4.43) by

$$\dot{\underline{C}}_{j,k:l,h,h'}^{\nu,\nabla} = (\dot{C}_{j,k:l,h,h'}^{\nu,\nabla} \vee 0) \wedge 1, \quad (4.44)$$

$$\dot{\underline{C}}_{j,k:l,h,h'}^{\nu,\Delta} = (\dot{C}_{j,k:l,h,h'}^{\nu,\Delta} \vee 0) \wedge 1. \quad (4.45)$$

*Remark 4.5.2.* As discussed in Remark 4.5.1, the smoothing distributions  $\nu_{\mathbf{u}}^{\mathcal{X}_{k:l}}$ ,  $\mathbf{u} \in [0, 1]^d$ , can be chosen such that  $C_{k:l}^{\nu}$  is a genuine copula under Condition 4.2.3. In that case, using the fact that  $C_{k:l}^{\nu}$  is a multivariate d.f. with standard uniform margins, we immediately obtain (see, e.g., Durante and Sempi, 2015, Lemma 1.2.14) that, for any  $\mathbf{u} \in [0, 1]^d$  and  $h, h'$  in  $[0, 1/2]$  such that  $h + h' > 0$ ,

$$C_{k:l}^{\nu}\{(\mathbf{u} + h\mathbf{e}_j) \wedge \mathbf{1}\} - C_{k:l}^{\nu}\{(\mathbf{u} - h'\mathbf{e}_j) \vee \mathbf{0}\} \leq (u_j + h) \wedge 1 - (u_j - h') \vee 0,$$

which implies that  $0 \leq \dot{C}_{j,k:l,h,h'}^{\nu,\Delta} \leq 1$  and thus that truncation of  $\dot{C}_{j,k:l,h,h'}^{\nu,\Delta}$  is not necessary in that case since  $\dot{\underline{C}}_{j,k:l,h,h'}^{\nu,\Delta}$  in (4.45) is equal to  $\dot{C}_{j,k:l,h,h'}^{\nu,\Delta}$ .

By analogy with (4.3), a second natural approach to obtain smooth partial derivative estimators consists of directly smoothing (4.38) and (4.39) and leads to the estimators

$$\dot{C}_{j,k:l,h,h'}^{\nabla,\nu}(\mathbf{u}) = \int_{[0,1]^d} \dot{C}_{j,k:l,h,h'}^{\nabla}(\mathbf{w}) d\nu_{\mathbf{u}^{k:l}}(\mathbf{w}), \quad \mathbf{u} \in [0,1]^d, \quad (4.46)$$

$$\dot{C}_{j,k:l,h,h'}^{\Delta,\nu}(\mathbf{u}) = \int_{[0,1]^d} \dot{C}_{j,k:l,h,h'}^{\Delta}(\mathbf{w}) d\nu_{\mathbf{u}^{k:l}}(\mathbf{w}), \quad \mathbf{u} \in [0,1]^d. \quad (4.47)$$

This time the order of the symbols  $\nu$  and  $\nabla$  (resp.  $\Delta$ ) is reversed indicating that it is the finite-differences-based estimator  $\dot{C}_{j,k:l,h,h'}^{\nabla}$  in (4.38) (resp.  $\dot{C}_{j,k:l,h,h'}^{\Delta}$  in (4.39)) that is smoothed. Versions of these estimators that necessarily take their values in  $[0,1]$  can be obtained by constructing them from the truncated estimators (4.40) and (4.41) instead, leading respectively to

$$\underline{\dot{C}}_{j,k:l,h,h'}^{\nabla,\nu}(\mathbf{u}) = \int_{[0,1]^d} \underline{\dot{C}}_{j,k:l,h,h'}^{\nabla}(\mathbf{w}) d\nu_{\mathbf{u}^{k:l}}(\mathbf{w}), \quad \mathbf{u} \in [0,1]^d, \quad (4.48)$$

$$\underline{\dot{C}}_{j,k:l,h,h'}^{\Delta,\nu}(\mathbf{u}) = \int_{[0,1]^d} \underline{\dot{C}}_{j,k:l,h,h'}^{\Delta}(\mathbf{w}) d\nu_{\mathbf{u}^{k:l}}(\mathbf{w}), \quad \mathbf{u} \in [0,1]^d. \quad (4.49)$$

Note that a third approach to obtain a smooth estimator of the  $j$ th partial derivative  $\dot{C}_j$  would consist of attempting to directly differentiate  $C_{k:l}^{\nu}$  in (4.3) with respect to its  $j$ th argument (provided of course that  $C_{k:l}^{\nu}$  is differentiable). The resulting estimator

$$\dot{C}_{j,k:l}^{\nu} = \frac{\partial C_{k:l}^{\nu}}{\partial u_j}$$

may exist only on the set  $V_j$  defined in Condition 4.2.8. This is the path followed by Janssen et al. (2016), who, for some integer  $m \geq 2$ , started from the empirical Bernstein copula  $C_{k:l,m}^{\text{Bern}}$  in (4.4) (which, as discussed in Remark 4.2.2, is a particular case of  $C_{k:l}^{\nu}$  in (4.3)). Let  $\dot{C}_{j,k:l,m}^{\text{Bern}} = \partial C_{k:l,m}^{\text{Bern}} / \partial u_j$  be the resulting estimator. Interestingly enough, from Lemma 4.7.8 in Appendix 4.7.5, we have that

$$\dot{C}_{j,k:l,m}^{\text{Bern}}(\mathbf{u}) = \int_{[0,1]^d} \dot{C}_{j,k:l,\frac{1}{m},0}^{\nabla}(\mathbf{w}) d\tilde{\mu}_{j,m,\mathbf{u}}(\mathbf{w}), \quad \mathbf{u} \in V_j, \quad (4.50)$$

where  $\dot{C}_{j,k:l,\frac{1}{m},0}^{\nabla}$  is given by (4.38) with  $h = 1/m$  and  $h' = 0$  and, for any  $\mathbf{u} \in [0,1]^d$ ,  $\tilde{\mu}_{j,m,\mathbf{u}}$  is the law of the random vector  $(\tilde{S}_{m,1,u_1}/m, \dots, \tilde{S}_{m,d,u_d}/m)$  whose components are independent such that, for  $i \in \{1, \dots, d\} \setminus \{j\}$ ,  $\tilde{S}_{m,i,u_i}$  is Binomial( $m, u_i$ ) while  $\tilde{S}_{m,j,u_j}$  is Binomial( $m - 1, u_j$ ). In other words, differentiating directly the empirical Bernstein copula  $C_{k:l,m}^{\text{Bern}}$  in (4.4) with respect to its  $j$ th argument leads to a special case of the estimator in (4.46). Notice that, since the measures  $\tilde{\mu}_{j,m,\mathbf{u}}$  are well-defined for any  $\mathbf{u} \in [0,1]^d$ , the integral in (4.50) is actually well-defined for any  $\mathbf{u} \in [0,1]^d$ . Hence, as we continue, we take (4.50) with  $\mathbf{u} \in [0,1]^d$  as the definition of  $\dot{C}_{j,k:l,m}^{\text{Bern}}$ .

The following result, proven in Appendix 4.7.5, shows that  $\dot{C}_{j,k:l,m}^{\text{Bern}}$  can be easily computed.

**Proposition 4.5.3.** *Given a stretch  $\mathcal{X}_{k:l}$ ,  $1 \leq k \leq l \leq n$ , of observations, we have that, for any  $j \in \{1, \dots, d\}$ ,  $\mathbf{u} \in [0, 1]^d$  and integer  $m \geq 2$ ,*

$$\begin{aligned} \dot{C}_{j,k:l,m}^{\text{Bern}}(\mathbf{u}) &= \frac{m}{l-k+1} \sum_{i=k}^l b_{m-1, u_j} \left\{ \lceil mR_{ij}^{k:l} / (l-k+1) \rceil - 1 \right\} \\ &\quad \times \prod_{\substack{t=1 \\ t \neq j}}^d \bar{B}_{m, u_t} \left\{ \lceil mR_{it}^{k:l} / (l-k+1) \rceil - 1 \right\}, \end{aligned} \quad (4.51)$$

where  $\lceil \cdot \rceil$  denotes the ceiling function and, for any  $p \in \mathbb{N}$  and  $u \in [0, 1]$ ,  $\bar{B}_{p,u}$  (resp.  $b_{p,u}$ ) is the survival (resp. probability mass) function of the Binomial( $p, u$ ).

### 4.5.3 Weak consistency

In order to study the weak consistency of the estimators of the partial derivatives of  $C$  defined in the previous subsection, it is necessary to link the bandwidth parameters in their expressions to the data (or, at least, to the amount of data) from which these estimators are computed. As we continue, for any  $n \in \mathbb{N}$  and any potential  $d$ -dimensional data set  $\mathbf{x} \in (\mathbb{R}^d)^n$ ,  $h(\mathbf{x})$  and  $h'(\mathbf{x})$  will denote the values of the left and right bandwidths for the data set  $\mathbf{x}$ . With this in mind, in the rest of this subsection, for the sake of a more compact notation, we shall write  $\dot{C}_{j,k:l}^{\nu, \nabla}$  (resp.  $\dot{C}_{j,k:l}^{\nu, \Delta}$ ,  $\dot{C}_{j,k:l}^{\nabla, \nu}$ , ...) for  $\dot{C}_{j,k:l,h,h'}^{\nu, \nabla}$  (resp.  $\dot{C}_{j,k:l,h,h'}^{\nu, \Delta}$ ,  $\dot{C}_{j,k:l,h,h'}^{\nabla, \nu}$ , ...) with the understanding that  $h = h(\mathcal{X}_{k:l})$  and  $h' = h'(\mathcal{X}_{k:l})$  are random variables. We impose in addition the following condition on the bandwidths.

**Condition 4.5.4** (Bandwidth condition). There exists positive sequences  $b_n \downarrow 0$  and  $b'_n \downarrow 0$  and constants  $L_2 \geq L_1 > 0$  such that, for all  $n \in \mathbb{N}$ ,  $b_n + b'_n \geq n^{-1/2}$ , and, for any  $\mathbf{x} \in (\mathbb{R}^d)^n$ ,  $L_1 b_n \leq h(\mathbf{x}) \leq (L_2 b_n) \wedge 1/2$  and  $L_1 b'_n \leq h'(\mathbf{x}) \leq (L_2 b'_n) \wedge 1/2$ .

As we shall see in Section 4.5.4, one meaningful possibility among many others is to consider that, for any  $n \in \mathbb{N}$  and  $\mathbf{x} \in \mathbb{R}^d$ , the left and right bandwidths for the data set  $\mathbf{x}$  are defined by  $h(\mathbf{x}) = h'(\mathbf{x}) = [M_2\{1 - |\tau(\mathbf{x})|\}^a + M_1]n^{-1/2} \wedge 1/2$ , where  $M_1, M_2 > 0$  are constants,  $\tau(\mathbf{x}) \in [-1, 1]$  is the value of the sample version of a suitable multivariate extension of Kendall's tau for the data set  $\mathbf{x}$  and  $a \in (0, \infty)$  is a fixed power. Roughly speaking, the bandwidths will be larger (resp. smaller) in the case of weakly (resp. strongly) cross-sectionally dependent data. It is easy to verify that Condition 4.5.4 holds for the previous definitions.

The following result, proven in Appendix 4.7.5, establishes the weak consistency of the smooth estimators of the first class in a sequential setting.

**Proposition 4.5.5** (Weak consistency in a sequential setting for the first class of smooth estimators). *Under Conditions 4.2.8, 4.2.9, 4.2.11 and 4.5.4, for any  $j \in \{1, \dots, d\}$ ,  $\delta \in (0, 1)$  and  $\varepsilon \in (0, 1/2)$ ,*

$$\sup_{\substack{(s,t) \in \Lambda \\ t-s \geq \delta}} \sup_{\substack{\mathbf{u} \in [0,1]^d \\ u_j \in [\varepsilon, 1-\varepsilon]}} \left| \dot{C}_{j, [ns]+1: [nt]}^{\nu, \Delta}(\mathbf{u}) - \dot{C}_j(\mathbf{u}) \right| = o_{\mathbb{P}}(1), \quad (4.52)$$

where  $\dot{C}_{j,k:l}^{\nu,\Delta}$  is defined in (4.43), and similarly for  $\dot{C}_{j,k:l}^{\nu,\nabla}$  in (4.42),  $\underline{\dot{C}}_{j,k:l}^{\nu,\nabla}$  in (4.44) and  $\underline{\dot{C}}_{j,k:l}^{\nu,\Delta}$  in (4.45).

*Remark 4.5.6.* An inspection of the proof of the previous result reveals that the second supremum in (4.52) can be replaced by a supremum over  $\mathbf{u} \in [0, 1]^d$  if  $\dot{C}_j$  happens to be continuous on  $[0, 1]^d$  instead of only satisfying Condition 4.2.8; see also Kojadinovic et al. (2011).

As a consequence of the previous proposition, we have that, under the conditions of Proposition 4.5.5, the estimators  $\underline{\dot{C}}_{j,k:l}^{\nu,\nabla}$  and  $\underline{\dot{C}}_{j,k:l}^{\nu,\Delta}$  satisfy Condition 4.4.1 since they are bounded in absolute value (by one) by construction.

The next result is an immediate corollary of Proposition 4.5.5 since the estimator in (4.38) (resp. (4.39)) is a particular case of the one in (4.42) (resp. (4.43)) when the smoothing distributions  $\nu_{\mathbf{u}}^{\mathbf{x},k:l}$ ,  $\mathbf{u} \in [0, 1]^d$ , in (4.3) are chosen to be Dirac measures at  $\mathbf{u} \in [0, 1]^d$  (the latter clearly satisfy Condition 4.2.9).

**Corollary 4.5.7** (Weak consistency in a sequential setting for the non-smooth finite-differences-based estimators). *Under Conditions 4.2.8, 4.2.11 and 4.5.4, for any  $j \in \{1, \dots, d\}$ ,  $\delta \in (0, 1)$  and  $\varepsilon \in (0, 1/2)$ ,*

$$\sup_{\substack{(s,t) \in \Lambda \\ t-s \geq \delta}} \sup_{\substack{\mathbf{u} \in [0,1]^d \\ u_j \in [\varepsilon, 1-\varepsilon]}} \left| \dot{C}_{j, [ns]+1: [nt]}^{\Delta}(\mathbf{u}) - \dot{C}_j(\mathbf{u}) \right| = o_{\mathbb{P}}(1),$$

where  $\dot{C}_{j,k:l}^{\Delta}$  is defined in (4.39), and similarly for  $\dot{C}_{j,k:l}^{\nabla}$  in (4.38),  $\underline{\dot{C}}_{j,k:l}^{\nabla}$  in (4.40) and  $\underline{\dot{C}}_{j,k:l}^{\Delta}$  in (4.41).

We now move to the second class of smooth estimators of the partial derivatives. As we shall see below, to establish their weak consistency, it suffices, among others, that the underlying smoothing distributions satisfy the following weaker version of Condition 4.2.9.

**Condition 4.5.8** (Weak variance condition). There exists a positive sequence  $a_n \downarrow 0$  such that, for any  $n \in \mathbb{N}$ ,  $\mathbf{x} \in (\mathbb{R}^d)^n$ ,  $\mathbf{u} \in [0, 1]^d$  and  $j \in \{1, \dots, d\}$ ,  $\text{Var}(W_{j,u_j}^{\mathbf{x}}) \leq a_n$ .

The following result is proven in Appendix 4.7.5.

**Proposition 4.5.9** (Weak consistency in a sequential setting for the second class of smooth estimators). *Under Conditions 4.2.8, 4.2.11, 4.5.4 and 4.5.8, for any  $j \in \{1, \dots, d\}$ ,  $\delta \in (0, 1)$  and  $\varepsilon \in (0, 1/2)$ ,*

$$\sup_{\substack{(s,t) \in \Lambda \\ t-s \geq \delta}} \sup_{\substack{\mathbf{u} \in [0,1]^d \\ u_j \in [\varepsilon, 1-\varepsilon]}} \left| \dot{C}_{j, [ns]+1: [nt]}^{\Delta, \nu}(\mathbf{u}) - \dot{C}_j(\mathbf{u}) \right| = o_{\mathbb{P}}(1), \quad (4.53)$$

where  $\dot{C}_{j,k:l}^{\Delta, \nu}$  is defined in (4.49), and similarly for  $\dot{C}_{j,k:l}^{\nabla, \nu}$  in (4.48).

One may wonder why the estimators  $\dot{C}_{j,k:l}^{\nabla,\nu}$  in (4.46) and  $\dot{C}_{j,k:l}^{\Delta,\nu}$  in (4.47) are not included in the previous proposition. Actually, upon additionally imposing that the left and right bandwidths of  $\dot{C}_{j,k:l}^{\nabla}$  in (4.38) and  $\dot{C}_{j,k:l}^{\Delta}$  in (4.39) (which are to be smoothed) are equal and in the absence of ties (see Condition 4.2.3), weak consistency can also be proven for the estimators  $\dot{C}_{j,k:l}^{\nabla,\nu}$  and  $\dot{C}_{j,k:l}^{\Delta,\nu}$  using the same technique of proof. An inspection of the proof and some additional thinking reveals that this follows from the fact that these estimators are bounded on  $[0, 1]^d$  in this case. When one of the bandwidths is zero, this is not necessarily the case anymore. This is also why the previous proposition cannot be directly used to establish the weak consistency of the Bernstein estimator in (4.50). For this estimator, one additionally needs to rely on the fact that the finite difference-based estimator that is smoothed is bounded on the support of the smoothing distributions. This is used in the proof in Appendix 4.7.5 of the next proposition.

**Proposition 4.5.10** (Weak consistency of the Bernstein estimator in a sequential setting). *Assume that Conditions 4.2.3, 4.2.8 and 4.2.11 hold and, for any  $i \in \mathbb{N}$ , let  $m_i = \lfloor Li^\theta \rfloor \vee 2$  for some constants  $L > 0$  and  $\theta \in (0, 1/2]$ . Then, for any  $j \in \{1, \dots, d\}$ ,  $\delta \in (0, 1)$  and  $\varepsilon \in (0, 1/2)$ ,*

$$\sup_{\substack{(s,t) \in \Lambda \\ t-s \geq \delta}} \sup_{\substack{\mathbf{u} \in [0,1]^d \\ u_j \in [\varepsilon, 1-\varepsilon]}} \left| \dot{C}_{j, \lfloor ns \rfloor + 1: \lfloor nt \rfloor, m_{\lfloor nt \rfloor - \lfloor ns \rfloor}}^{\text{Bern}} - \dot{C}_j(\mathbf{u}) \right| = o_{\mathbb{P}}(1), \quad (4.54)$$

where  $\dot{C}_{j,k:l,m}^{\text{Bern}}$  is defined in (4.50). In addition, for any  $n \in \mathbb{N}$ ,

$$\sup_{(s,t,\mathbf{u}) \in \Lambda \times [0,1]^d} \left| \dot{C}_{j, \lfloor ns \rfloor + 1: \lfloor nt \rfloor, m_{\lfloor nt \rfloor - \lfloor ns \rfloor}}^{\text{Bern}}(\mathbf{u}) \right| \leq 1 + L \vee 2. \quad (4.55)$$

Note that our technique of proof allows us to establish uniform weak consistency of the estimator even for  $\theta = 1/2$ , whereas the approach used in the proof of Proposition 3.1 of Bouezmarni et al. (2013) (for Bernstein copula density estimators and which could be adapted to first-order partial derivative estimators) leads to the result only for  $\theta < 1/2$ . Another consequence of the previous result is that the Bernstein partial derivative estimator as parametrized in Proposition 4.5.10 can satisfy Condition 4.4.1.

#### 4.5.4 Finite-sample performance of selected estimators

The aim of this subsection is to compare the finite-sample performance of some of the estimators introduced previously. Specifically, for each  $n \in \{10, 20, \dots, 100\}$ , each data generating copula  $C$  and each partial derivative estimator  $\dot{C}_{j,1:n}$  under investigation, we estimated its integrated mean squared error

$$\text{IMSE}(\dot{C}_{j,1:n}) = \int_{[0,1]^d} \mathbb{E} \left[ \left\{ \dot{C}_{j,1:n}(\mathbf{u}) - \dot{C}_j(\mathbf{u}) \right\}^2 \right] d\mathbf{u}.$$

To do so, we applied the trick described in detail in Appendix B of Segers et al. (2017) allowing to compute  $\text{IMSE}(\hat{C}_{j,1:n})$  as a single expectation and proceeded by Monte Carlo simulation using 20 000 independent random samples of size  $n$  from  $C$ .

In a first experiment, we compared estimators of the form  $\hat{C}_{j,1:n,h,h'}^{\nu,\nabla}$  in (4.44) to estimators of the form  $\hat{C}_{j,1:n,h,h'}^{\nu,\Delta}$  in (4.45) for deterministic bandwidths  $h = h' = n^{-1/2} \wedge 1/2$ . Specifically, we considered estimators based, respectively, on finite-differences of the classical empirical copula  $C_{1:n}$  in (4.2), on finite-differences of the empirical beta copula  $C_{1:n}^{\text{Bin}}$  defined in (4.17), and on finite-differences of its data-adaptive extension  $C_{1:n}^{\text{BetaB4}}$  defined at the end of Section 4.2.1. As data-generating copula  $C$ , we considered the bivariate or trivariate Clayton or Gumbel–Hougaard copula with bivariate margins with a Kendall’s tau of  $\tau \in \{0, 0.25, 0.5, 0.75\}$  as well as the bivariate Frank copula with a Kendall’s tau of  $\tau \in \{0, -0.25, -0.5, -0.75\}$ . Note that, since all data-generating copulas are exchangeable, it suffices to focus on only one partial derivative estimator, say the first one. As expected, the integrated mean squared error of estimators of the form  $\hat{C}_{j,1:n,h,h'}^{\nu,\Delta}$  was always found to be (substantially) below that of the corresponding estimator  $\hat{C}_{j,1:n,h,h'}^{\nu,\nabla}$ , confirming that the adjusted numerator in (4.43) compared to the one in (4.42) helps indeed to improve the finite-sample performance of finite-difference-based estimators.

In a second experiment, we compared the aforementioned three estimators of the form  $\hat{C}_{j,1:n,h,h'}^{\nu,\Delta}$  in (4.45). They will be denoted as  $\hat{C}_{j,1:n,h,h'}^{\Delta} = \hat{C}_{j,1:n,h,h'}^{\text{Dirac},\Delta}$ ,  $\hat{C}_{j,1:n,h,h'}^{\text{Bin},\Delta}$  and  $\hat{C}_{j,1:n,h,h'}^{\text{BetaB4},\Delta}$  as we continue. As could have been expected from the experiments of Kojadinovic and Yi (2022) comparing the underlying copula estimators, it is the estimator  $\hat{C}_{j,1:n,h,h'}^{\text{BetaB4},\Delta}$  that always displayed the lowest integrated mean squared error, followed by  $\hat{C}_{j,1:n,h,h'}^{\text{Bin},\Delta}$  and  $\hat{C}_{j,1:n,h,h'}^{\Delta}$ .

We next investigated the influence of the bandwidths on the integrated mean squared error of  $\hat{C}_{j,1:n,h,h'}^{\text{BetaB4},\Delta}$ . Deterministic bandwidths of the form  $h = h' = (Ln^{-1/2}) \wedge 1/2$  were considered with  $L \in \{0.5, 1, 2, 4\}$ . The corresponding integrated mean squared errors are represented in the first column of graphs of Figure 4.3 (resp. Figure 4.4) when the data-generating copula is the bivariate Frank copula with negative dependence (resp. the trivariate Gumbel–Hougaard copula). The legend “BetaB 0.5” refers to the estimator with  $L = 0.5$  and so on. As one can see, the weaker the cross-sectional dependence, the larger the (constant  $L$  in the expression of the) bandwidths should be.

An inspection of (4.46) and (4.47) and some thinking reveals that estimators from the second class can be difficult to compute in practice. For that reason, in our experiments, we solely focused on the Bernstein estimator  $\hat{C}_{j,1:n,m}^{\text{Bern}}$  in (4.50) which can be computed using (4.51). Mimicking the previous experiment, we considered a deterministic choice for the parameter  $m$  of the form  $m = \lfloor Ln^{1/2} \rfloor \vee 2$  with  $L \in \{0.5, 1, 2, 4\}$ . The corresponding integrated mean squared errors are represented in the second column of graphs of Figure 4.3 (resp. Figure 4.4) when the data-generating copula is the bivariate Frank copula with negative dependence (resp. the trivariate Gumbel–Hougaard copula). The legend “Bern 0.5” refers to the estimator with  $L = 0.5$  and so on. As one can see, this time, the stronger the cross-sectional dependence, the

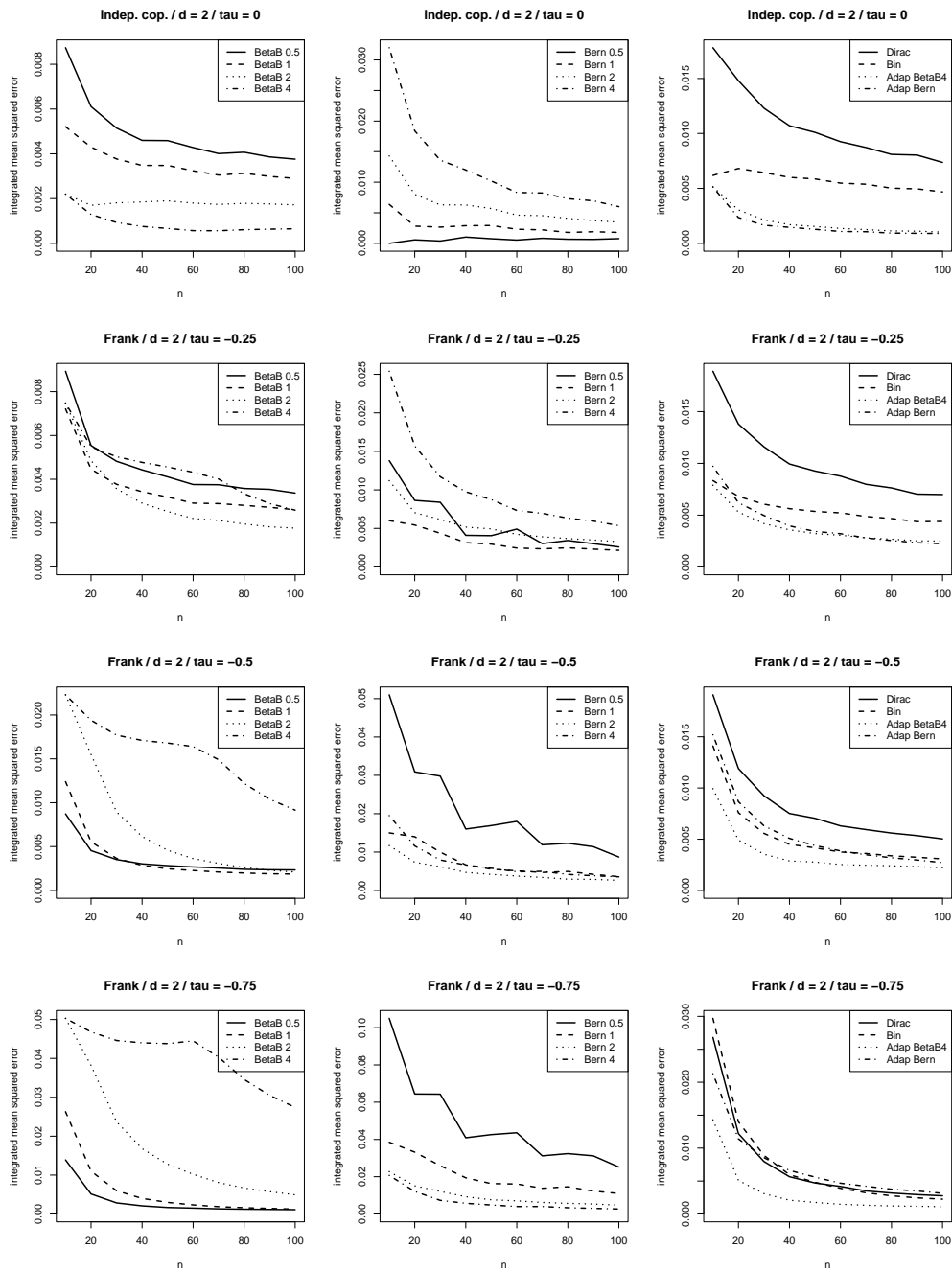


FIGURE 4.3: Estimated integrated mean squared errors against  $n \in \{10, 20, \dots, 100\}$  of four estimators of  $\hat{C}_1$  when  $C$  is the bivariate Frank copula with a Kendall's tau in  $\{0, -0.25, -0.5, -0.75\}$ .

larger the (constant  $L$  in the expression of the) parameter  $m$  should be. This is of course not surprising given the previous experiment and since  $1/m$  plays the role of a bandwidth.

The two previous experiments suggested to focus on data-adaptive bandwidths for the estimators  $\hat{C}_{j,1:n,h,h'}^{\text{BetaB4},\Delta}$  and  $\hat{C}_{j,1:n,m}^{\text{Bern}}$ . Specifically, for  $d \in \{2, 3\}$  and a data set  $\mathbf{x} \in (\mathbb{R}^d)^n$ , we considered the settings  $h(\mathbf{x}) = h'(\mathbf{x}) = ([4\{1 - |\tau(\mathbf{x})|\}]^6 + 1/2)n^{-1/2}) \wedge 1/2$  for  $\hat{C}_{j,1:n,h,h'}^{\text{BetaB4},\Delta}$  and  $m(\mathbf{x}) =$

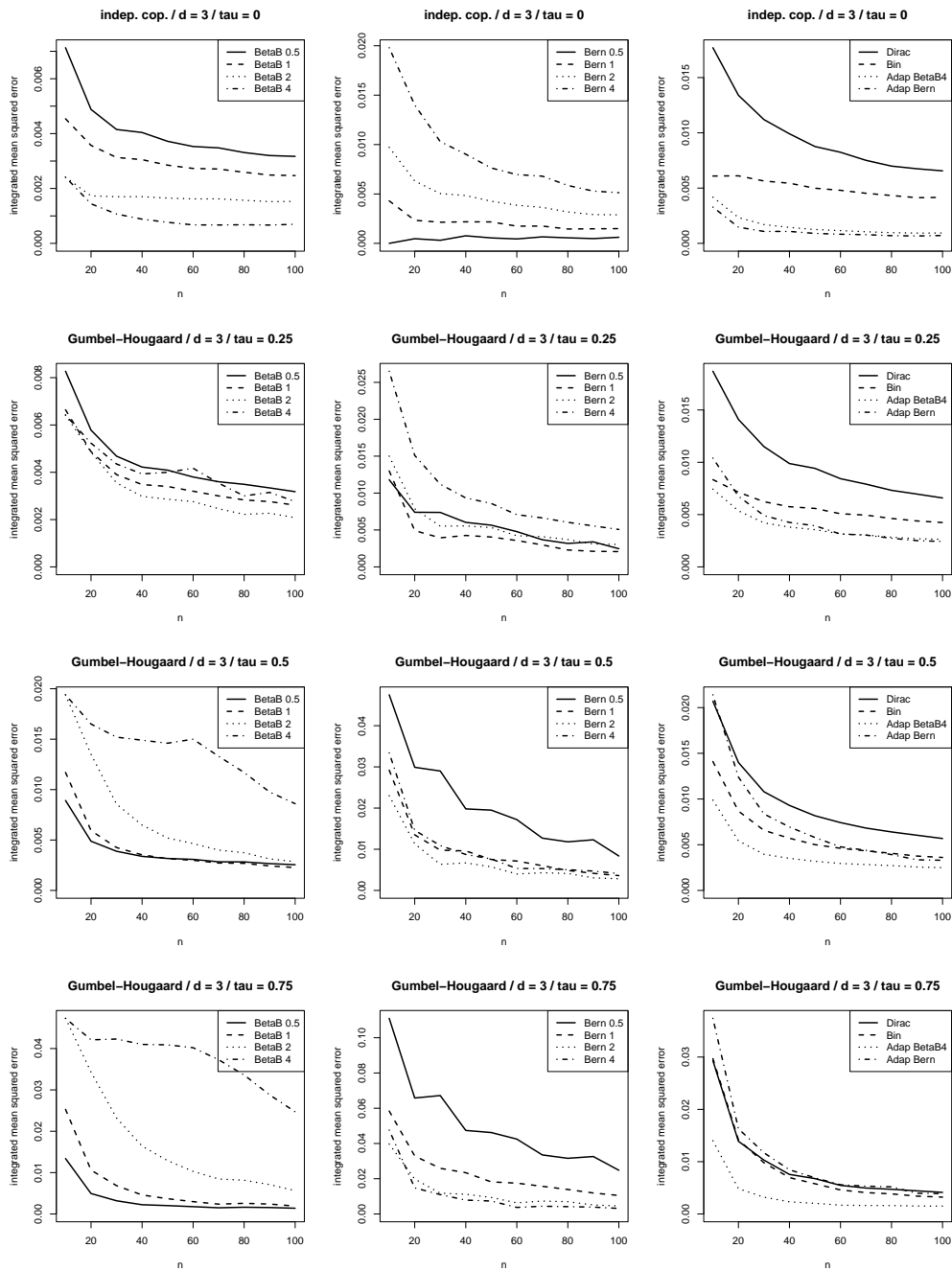


FIGURE 4.4: Estimated integrated mean squared errors against  $n \in \{10, 20, \dots, 100\}$  of four estimators of  $\hat{C}_1$  when  $C$  is the trivariate Gumbel–Hougaard copula whose bivariate margins have a Kendall’s tau in  $\{0, 0.25, 0.5, 0.75\}$ .

$[\{4|\tau(\mathbf{x})|^{3/2} + 1/2\}n^{1/2}] \vee 2$  for  $\hat{C}_{j,1:n,m}^{\text{Bern}}$ , where  $\tau(\mathbf{x}) \in [-1, 1]$  is the average of the values of the sample version of Kendall’s tau computed from the bivariate margins of the data set  $\mathbf{x}$ . The integrated mean squared errors of the resulting data-adaptive estimators are represented in the third column of graphs of Figure 4.3 (resp. Figure 4.4) when the data-generating copula is the bivariate Frank copula with negative dependence (resp. the trivariate Gumbel–Hougaard copula). The legends “Adap BetaB4” and “Adap Bern” refer to the above-mentioned versions



of the estimators  $\hat{C}_{j,1:n,h,h'}^{\text{BetaB4},\Delta}$  and  $\hat{C}_{j,1:n,m}^{\text{Bern}}$ , respectively, while “Dirac” and “Bin” refer to the benchmark estimators  $\hat{C}_{j,1:n,h,h'}^{\Delta} = \hat{C}_{j,1:n,h,h'}^{\text{Dirac},\Delta}$  and  $\hat{C}_{j,1:n,h,h'}^{\text{Bin},\Delta}$  with deterministic bandwidths  $h = h' = n^{-1/2} \wedge 1/2$  based on the empirical copula and the empirical beta copula, respectively. Overall, it is the data-adaptive estimator  $\hat{C}_{j,1:n,h,h'}^{\text{BetaB4},\Delta}$  that displays the best finite-sample behavior. The data-adaptive Bernstein estimator  $\hat{C}_{j,1:n,m}^{\text{Bern}}$  appears to be competitive only when the data-generating copula is close to the independence copula.

## 4.6 Conclusion

Smooth nonparametric copula estimators, such as the empirical beta copula proposed by Segers et al. (2017) or its data-adaptive extension studied in Kojadinovic and Yi (2022), can be substantially better estimators than the classical empirical copula in small samples. To use such estimators in inference procedures, one typically needs to rely on resampling techniques.

As investigated in Section 4.3, in the case of i.i.d. observations, a smooth bootstrap *à la* Kiriliouk et al. (2021) can be asymptotically valid for a large class of smooth estimators that can be expressed as mixtures of d.f.s. When based on the empirical beta copula, Kiriliouk et al. (2021) found such a smooth bootstrap to be a competitive alternative to the multiplier bootstrap while being substantially simpler to implement. An empirical finding of this work is that the smooth bootstrap based on the data-adaptive extension of the empirical beta copula proposed in Kojadinovic and Yi (2022) seems to lead to even better-behaved inference procedures than the former as it copes better with stronger dependence.

Unfortunately, such smooth bootstraps cannot be used anymore in the time series setting. A second contribution of this work was to study both theoretically and empirically smooth extensions of the sequential dependent multiplier bootstrap of Bücher and Kojadinovic (2016). As illustrated at the end of the fourth section, the latter can for instance be used to derive smooth change-point detection tests which are likely to be more sensitive to early or late changes than their non-smooth counterparts since, as already mentioned, smooth estimators are likely to be more accurate than the empirical copula when computed from small subsets of observations.

In connection with the multiplier bootstrap, a third contribution of this work was the study of the weak consistency and finite-sample performance of two classes of smooth estimators of the first-order partial derivatives of the copula. The obtained results may be of independent interest since such estimators have applications in mean and quantile regression as they lead to estimators of the conditional distribution function. From an empirical perspective, our investigations led to the proposal of a smooth data-adaptive estimator of the first-order partial derivatives of the copula that substantially outperforms, among others, the Bernstein estimator studied in Janssen et al. (2016).

## 4.7 Proofs

### 4.7.1 Proof of Corollary 4.2.13

The proof of Corollary 4.2.13 is based on the following two lemmas.

**Lemma 4.7.1.** *Let  $\mathbb{X}_n$  be a process in  $\ell^\infty(\Lambda \times [0, 1]^d)$  such that for all  $\mathbf{u} \in [0, 1]^d$  and  $s \in [0, 1]$ ,  $\mathbb{X}_n(s, s, \mathbf{u}) = 0$ . Furthermore, assume that  $\mathbb{X}_n \rightsquigarrow \mathbb{X}$  in  $\ell^\infty(\Lambda \times [0, 1]^d)$  where  $\mathbb{X}$  has continuous trajectories almost surely. Then, under Condition 4.5.8 (which is implied by Condition 4.2.9),*

$$\sup_{(s,t,\mathbf{u}) \in \Lambda \times [0,1]^d} \left| \int_{[0,1]^d} \mathbb{X}_n(s, t, \mathbf{w}) d\nu_{\mathbf{u}}^{\mathbf{x}^{[ns]+1:[nt]}}(\mathbf{w}) - \mathbb{X}_n(s, t, \mathbf{u}) \right| = o_{\mathbb{P}}(1), \quad (4.56)$$

$$\sup_{(s,t,\mathbf{u}) \in \Lambda \times [0,1]^d} \left| \int_{[0,1]^d} \mathbb{X}_n(s, t, \mathbf{w}) d\nu_{\mathbf{u}}^{\mathbf{x}^{1:n}}(\mathbf{w}) - \mathbb{X}_n(s, t, \mathbf{u}) \right| = o_{\mathbb{P}}(1). \quad (4.57)$$

*Proof.* The first claim was proven in the proof of Lemma 32 of [Kojadinovic and Yi \(2022\)](#). The proof of (4.57) is very similar.  $\blacksquare$

**Lemma 4.7.2.** *Assume that Conditions 4.2.8 and 4.2.9 hold. Then, almost surely,*

$$\sup_{(s,t,\mathbf{u}) \in \Lambda \times [0,1]^d} \sqrt{n} \lambda_n(s, t) \left| \int_{[0,1]^d} C(\mathbf{w}) d\nu_{\mathbf{u}}^{\mathbf{x}^{[ns]+1:[nt]}}(\mathbf{w}) - C(\mathbf{u}) \right| = o(1), \quad (4.58)$$

$$\sup_{(s,t,\mathbf{u}) \in \Lambda \times [0,1]^d} \sqrt{n} \lambda_n(s, t) \left| \int_{[0,1]^d} C(\mathbf{w}) d\nu_{\mathbf{u}}^{\mathbf{x}^{1:n}}(\mathbf{w}) - C(\mathbf{u}) \right| = o(1). \quad (4.59)$$

*Proof.* The first claim was proven in the proof of Lemma 33 of [Kojadinovic and Yi \(2022\)](#). The proof of (4.59) is an immediate consequence of the fact that the left-hand side of (4.59) is almost surely smaller than

$$\sup_{\mathbf{u} \in [0,1]^d} \sqrt{n} \left| \int_{[0,1]^d} C(\mathbf{w}) d\nu_{\mathbf{u}}^{\mathbf{x}^{1:n}}(\mathbf{w}) - C(\mathbf{u}) \right|,$$

which is smaller than the left-hand side of (4.58) with probability 1.  $\blacksquare$

**Proof of Corollary 4.2.13.** Combining Theorem 4.2.10 and Theorem 4.2.12, by the triangle inequality, we immediately obtain that

$$\sup_{(s,t,\mathbf{u}) \in \Lambda \times [0,1]^d} \left| \mathbb{C}_n^\nu(s, t, \mathbf{u}) - \tilde{\mathbb{C}}_n(s, t, \mathbf{u}) \right| = o_{\mathbb{P}}(1),$$

where  $\tilde{\mathbb{C}}_n$  is defined in (4.13). It thus remains to show that

$$\begin{aligned} \sup_{(s,t,\mathbf{u}) \in \Lambda \times [0,1]^d} \left| \tilde{\mathbb{C}}_n(s, t, \mathbf{u}) - \tilde{\mathbb{C}}_n^\nu(s, t, \mathbf{u}) \right| &= o_{\mathbb{P}}(1), \\ \sup_{(s,t,\mathbf{u}) \in \Lambda \times [0,1]^d} \left| \tilde{\mathbb{C}}_n(s, t, \mathbf{u}) - \bar{\mathbb{C}}_n^\nu(s, t, \mathbf{u}) \right| &= o_{\mathbb{P}}(1). \end{aligned}$$

We only prove the first claim, the proof of the second one being similar. For any  $(s, t, \mathbf{u}) \in \Lambda \times [0, 1]^d$ , let

$$\check{\mathbb{B}}_n^\nu(s, t, \mathbf{u}) = \int_{[0,1]^d} \mathbb{B}_n(s, t, \mathbf{w}) d\nu_{\mathbf{u}}^{\mathbf{x}_{[ns]+1:[nt]}}(\mathbf{w}).$$

Under Condition 4.2.11,  $\mathbb{B}_n \rightsquigarrow \mathbb{B}_C$  in  $\ell^\infty(\Lambda \times [0, 1]^d)$ , where  $\mathbb{B}_C$  has continuous trajectories almost surely. We then obtain from (4.56) in Lemma 4.7.1 that

$$\sup_{(s,t,\mathbf{u}) \in \Lambda \times [0,1]^d} \left| \check{\mathbb{B}}_n^\nu(s, t, \mathbf{u}) - \mathbb{B}_n(s, t, \mathbf{u}) \right| = o_{\mathbb{P}}(1),$$

and, furthermore, since Conditions 4.2.8 and 4.2.9 hold, from (4.58) in Lemma 4.7.2, that

$$\sup_{(s,t,\mathbf{u}) \in \Lambda \times [0,1]^d} \left| \check{\mathbb{B}}_n^\nu(s, t, \mathbf{u}) - \tilde{\mathbb{B}}_n^\nu(s, t, \mathbf{u}) \right| = o(1),$$

with probability one, where  $\tilde{\mathbb{B}}_n^\nu$  is defined in (4.15), which implies that

$$\sup_{(s,t,\mathbf{u}) \in \Lambda \times [0,1]^d} \left| \tilde{\mathbb{B}}_n^\nu(s, t, \mathbf{u}) - \mathbb{B}_n(s, t, \mathbf{u}) \right| = o_{\mathbb{P}}(1). \quad (4.60)$$

Moreover, from the triangle inequality, we have

$$\begin{aligned} \sup_{\substack{(s,t) \in \Lambda \\ \mathbf{u} \in [0,1]^d}} \left| \tilde{\mathcal{C}}_n^\nu(s, t, \mathbf{u}) - \tilde{\mathcal{C}}_n(s, t, \mathbf{u}) \right| &\leq \sup_{\substack{(s,t) \in \Lambda \\ \mathbf{u} \in [0,1]^d}} \left| \tilde{\mathbb{B}}_n^\nu(s, t, \mathbf{u}) - \mathbb{B}_n(s, t, \mathbf{u}) \right| \\ &+ \sum_{j=1}^d \sup_{\mathbf{u} \in [0,1]^d} \left| \dot{C}_j(\mathbf{u}) \right| \sup_{\substack{(s,t) \in \Lambda \\ \mathbf{u} \in [0,1]^d}} \left| \tilde{\mathbb{B}}_n^\nu(s, t, \mathbf{u}^{(j)}) - \mathbb{B}_n(s, t, \mathbf{u}^{(j)}) \right|. \end{aligned}$$

The terms on the right-hand side of the previous display converge to zero in probability as a consequence of (4.60) and the fact that  $0 \leq \dot{C}_j \leq 1$ .  $\blacksquare$

## 4.7.2 Proofs of Proposition 4.3.3 and Lemma 4.7.4

**Proof of Proposition 4.3.3.** Fix  $n \in \mathbb{N}$ ,  $\mathbf{x} \in (\mathbb{R}^d)^n$  and  $\mathbf{r} \in \{1, \dots, n\}^d$  and let us check that  $\mathcal{K}_{\mathbf{r}}^{\mathbf{x}}$ , which can be expressed as in (4.7) under Condition 4.2.4, is a multivariate d.f.

By Condition 4.2.5, for any  $j \in \{1, \dots, d\}$ , the function  $\mathcal{K}_{r_j, j}^{\mathbf{x}}$  defined by  $\mathcal{K}_{r_j, j}^{\mathbf{x}}(u) = \bar{\mathcal{F}}_{j, u}^{\mathbf{x}}\{(r_j - 1)/n\}$ ,  $u \in [0, 1]$ , is a univariate d.f. on  $[0, 1]$ . Indeed,  $\mathcal{K}_{r_j, j}^{\mathbf{x}}$  is right-continuous and increasing on  $[0, 1]$  and, by properties of the smoothing distributions,

$$\begin{aligned} \mathcal{K}_{r_j, j}^{\mathbf{x}}(0) &= \bar{\mathcal{F}}_{j, 0}^{\mathbf{x}}\{(r_j - 1)/n\} = \mathbb{P}\{W_{j, 0}^{\mathbf{x}} > (r_j - 1)/n\} = \mathbb{P}\{0 > (r_j - 1)/n\} = 0, \\ \mathcal{K}_{r_j, j}^{\mathbf{x}}(1) &= \bar{\mathcal{F}}_{j, 1}^{\mathbf{x}}\{(r_j - 1)/n\} = \mathbb{P}\{W_{j, 1}^{\mathbf{x}} > (r_j - 1)/n\} = \mathbb{P}\{1 > (r_j - 1)/n\} = 1. \end{aligned}$$

Using additionally Condition 4.2.6, the expression of  $\mathcal{K}_r^x$  in (4.7) can then be further simplified to

$$\mathcal{K}_r^x(\mathbf{u}) = \bar{\mathcal{C}}^x \{ \mathcal{K}_{r_1,1}^x(u_1), \dots, \mathcal{K}_{r_d,d}^x(u_d) \}, \quad \mathbf{u} \in [0, 1]^d.$$

From Sklar's Theorem (Sklar, 1959),  $\mathcal{K}_r^x$  is thus a d.f. on  $[0, 1]^d$  with univariate margins  $\mathcal{K}_{r_1,1}^x, \dots, \mathcal{K}_{r_d,d}^x$  and copula  $\bar{\mathcal{C}}^x$ . ■

The proof of Lemma 4.7.4 below is based on the following lemma.

**Lemma 4.7.3.** *For any  $n \in \mathbb{N}$  and  $t \in [0, 1]$ , let  $\bar{B}_{n,t}$  be the survival function of a Binomial( $n, t$ ). Then, for any  $n \in \mathbb{N}$  and  $w \in [0, n)$ , the function  $t \mapsto \bar{B}_{n,t}(w)$  is strictly increasing on  $[0, 1]$ .*

*Proof.* Fix  $n \in \mathbb{N}$  and  $w \in [0, n)$ . Since  $t \mapsto \bar{B}_{n,t}(w)$  is continuous on  $[0, 1]$ , it suffices to prove that, for any  $t \in (0, 1)$ ,  $\frac{\partial}{\partial t} \{ \bar{B}_{n,t}(w) \} > 0$ . We have

$$\begin{aligned} \frac{\partial}{\partial t} \{ \bar{B}_{n,t}(w) \} &= \frac{\partial}{\partial t} \left\{ \sum_{k=\lfloor w \rfloor + 1}^n \binom{n}{k} t^k (1-t)^{n-k} \right\} = \frac{\partial}{\partial t} \left\{ \sum_{k=\lfloor w \rfloor + 1}^n \binom{n}{k} t^k (1-t)^{n-k} \right\} \\ &= \sum_{k=\lfloor w \rfloor + 1}^n \frac{n!}{k!(n-k)!} \{ k t^{k-1} (1-t)^{n-k} - (n-k) t^k (1-t)^{n-k-1} \} \\ &= \sum_{k=\lfloor w \rfloor + 1}^n \frac{n!}{k!(n-k)!} k t^{k-1} (1-t)^{n-k} - \sum_{k=\lfloor w \rfloor + 1}^{n-1} \frac{n!}{k!(n-k)!} (n-k) t^k (1-t)^{n-k-1} \\ &= \sum_{k=\lfloor w \rfloor + 1}^n \frac{n!}{(k-1)!(n-k)!} t^{k-1} (1-t)^{n-k} - \sum_{k=\lfloor w \rfloor + 1}^{n-1} \frac{n!}{k!(n-k-1)!} t^k (1-t)^{n-k-1} \\ &= \sum_{k=\lfloor w \rfloor + 1}^n \frac{n!}{(k-1)!(n-k)!} t^{k-1} (1-t)^{n-k} - \sum_{k=\lfloor w \rfloor + 2}^n \frac{n!}{(k-1)!(n-k)!} t^{k-1} (1-t)^{n-k} \\ &= \frac{n!}{\lfloor w \rfloor!(n-\lfloor w \rfloor-1)!} t^{\lfloor w \rfloor} (1-t)^{n-\lfloor w \rfloor-1} > 0. \end{aligned}$$

■

**Lemma 4.7.4.** *For any  $n \in \mathbb{N}$ ,  $t \in [0, 1]$  and  $\rho \in (1, n)$ , let  $\bar{B}_{n,t,\rho}$  be the survival function of a Beta-Binomial( $n, \alpha, \beta$ ), where  $\alpha = t(n-\rho)/(\rho-1)$  and  $\beta = (1-t)(n-\rho)/(\rho-1)$ . Then, for any  $n \in \mathbb{N}$ ,  $\rho \in (1, n)$  and  $w \in [0, n)$ , the function  $t \mapsto \bar{B}_{n,t,\rho}(w)$  is strictly increasing on  $[0, 1]$ .*

*Proof.* First, notice that, for any  $t \in [0, 1]$ ,  $\rho \in (1, n)$  and  $w \in [0, n)$ , by definition of the beta-binomial distribution,

$$\bar{B}_{n,t,\rho}(w) = \mathbb{E}_\Theta \{ \bar{B}_{n,\Theta}(w) \}, \tag{4.61}$$

where  $\bar{B}_{n,t}$  is the survival function of a Binomial( $n, t$ ) and  $\Theta$  is Beta( $\alpha, \beta$ ) with  $\alpha = t(n-\rho)/(\rho-1)$  and  $\beta = (1-t)(n-\rho)/(\rho-1)$ . From Lemma 30 in Kojadinovic and Yi (2022), we have that, for any  $n \in \mathbb{N}$ ,  $\rho \in (1, n)$  and  $w \in [0, n)$ , the function  $t \mapsto \bar{B}_{n,t,\rho}(w)$  is increasing on  $[0, 1]$ . It thus suffices to show strict increasingness. Let us prove this by contradiction. Suppose that

there exists  $0 \leq t_1 < t_2 \leq 1$  such that  $\bar{\mathcal{B}}_{n,t_1,\rho}(w) = \bar{\mathcal{B}}_{n,t_2,\rho}(w)$  for some  $n \in \mathbb{N}$ ,  $\rho \in (1, n)$  and  $w \in [0, n)$ . Then, from (4.61), we have that

$$\mathbb{E}_{\Theta_1}\{\bar{B}_{n,\Theta_1}(w)\} = \mathbb{E}_{\Theta_2}\{\bar{B}_{n,\Theta_2}(w)\}, \quad (4.62)$$

where  $\Theta_1$  (resp.  $\Theta_2$ ) is  $\text{Beta}(\alpha_1, \beta_1)$  (resp.  $\text{Beta}(\alpha_2, \beta_2)$ ) with  $\alpha_1 = t_1(n - \rho)/(\rho - 1)$  and  $\beta_1 = (1 - t_1)(n - \rho)/(\rho - 1)$  (resp.  $\alpha_2 = t_2(n - \rho)/(\rho - 1)$  and  $\beta_2 = (1 - t_2)(n - \rho)/(\rho - 1)$ ). From the proof of Lemma 29 in [Kojadinovic and Yi \(2022\)](#), we have that  $\Theta_1 \leq_{st} \Theta_2$ , where  $\leq_{st}$  denotes the usual stochastic order. Using additionally (4.62) and the fact that the function  $t \mapsto \bar{B}_{n,t}(w)$  is strictly increasing on  $[0, 1]$  from Lemma 4.7.3, we have, according to Theorem 1.A.8 of [Shaked and Shanthikumar \(2007\)](#), that  $\Theta_1$  and  $\Theta_2$  have the same distribution. This contradicts the fact that  $t_1 < t_2$ .  $\blacksquare$

### 4.7.3 Proof of Theorem 4.3.5

The proof of Theorem 4.3.5 is based on two lemmas which we show first.

Let  $\Phi$  be the map from  $\ell^\infty([0, 1]^d)$  to  $\ell^\infty([0, 1]^d)$  defined for any d.f  $H$  on  $[0, 1]^d$  whose univariate margins  $H_1, \dots, H_d$  do not assign mass at zero by

$$\Phi(H)(\mathbf{u}) = H\{H_1^{-1}(u_1), \dots, H_d^{-1}(u_d)\}, \quad \mathbf{u} \in [0, 1]^d. \quad (4.63)$$

**Lemma 4.7.5.** *Assume that the random vectors in  $\mathcal{X}_{1:n}$  are i.i.d. and that Condition 4.5.8 holds. Then, almost surely,*

$$\sup_{\mathbf{u} \in [0, 1]^d} |C_{1:n}^\nu(\mathbf{u}) - C(\mathbf{u})| = o(1), \quad (4.64)$$

where  $C_{1:n}^\nu$  is defined in (4.3).

*Proof.* The supremum on the left-hand side of (4.64) is smaller than  $I_n + J_n$ , where

$$I_n = \sup_{\mathbf{u} \in [0, 1]^d} \left| \int_{[0, 1]^d} C_{1:n}(\mathbf{w}) d\nu_{\mathbf{u}}^{\mathcal{X}_{1:n}}(\mathbf{w}) - \int_{[0, 1]^d} C(\mathbf{w}) d\nu_{\mathbf{u}}^{\mathcal{X}_{1:n}}(\mathbf{w}) \right|,$$

$$J_n = \sup_{\mathbf{u} \in [0, 1]^d} \left| \int_{[0, 1]^d} C(\mathbf{w}) d\nu_{\mathbf{u}}^{\mathcal{X}_{1:n}}(\mathbf{w}) - C(\mathbf{u}) \right|.$$

*Term  $I_n$ :* From the triangle inequality,  $I_n$  is smaller than

$$\sup_{\mathbf{u} \in [0, 1]^d} |C_{1:n}(\mathbf{u}) - C(\mathbf{u})| \leq I_n' + I_n'' + I_n''',$$

where

$$\begin{aligned} I'_n &= \sup_{\mathbf{u} \in [0,1]^d} |C_{1:n}(\mathbf{u}) - \Phi(G_{1:n})(\mathbf{u})|, & I''_n &= \sup_{\mathbf{u} \in [0,1]^d} |\Phi(G_{1:n})(\mathbf{u}) - G_{1:n}(\mathbf{u})|, \\ I'''_n &= \sup_{\mathbf{u} \in [0,1]^d} |G_{1:n}(\mathbf{u}) - C(\mathbf{u})|, \end{aligned}$$

where the map  $\Phi$  is defined in (4.63) and  $G_{1:n}$  is empirical d.f. of the unobservable random sample  $\mathbf{U}_1, \dots, \mathbf{U}_n$  obtained from  $\mathcal{X}_{1:n}$  by the probability integral transformations  $U_{ij} = F_j(X_{ij})$ ,  $i \in \{1, \dots, n\}$ ,  $j \in \{1, \dots, d\}$ . Using the well-known facts (see, e.g., Segers, 2012) that  $\Phi(G_{1:n}) = \Phi(F_{1:n})$ , where  $F_{1:n}$  is the empirical d.f. of  $\mathcal{X}_{1:n}$ , and

$$\sup_{\mathbf{u} \in [0,1]^d} |C_{1:n}(\mathbf{u}) - \Phi(F_{1:n})(\mathbf{u})| \leq \frac{d}{n},$$

we obtain that  $I'_n = o(1)$ . Furthermore, from the Glivenko-Cantelli lemma (see, e.g., van der Vaart, 1998, Theorem 19.1),  $I'''_n = o(1)$  with probability one. Finally, using a well-known property of multivariate d.f.s (see, e.g., Durante and Sempi, 2015, Lemma 1.2.14), the well-known fact, for any  $j \in \{1, \dots, d\}$ ,  $\sup_{u \in [0,1]} |G_{1:n,j}^{-1}(u) - u| = \sup_{u \in [0,1]} |G_{1:n,j}(u) - u|$  (see, e.g., Shorack and Wellner, 1986, Chapter 3) and, again, the Glivenko-Cantelli lemma, we obtain that, almost surely,

$$I''_n \leq \sum_{j=1}^d \sup_{u \in [0,1]} |G_{1:n,j}^{-1}(u) - u| = \sum_{j=1}^d \sup_{u \in [0,1]} |G_{1:n,j}(u) - u| = o(1).$$

*Term  $J_n$ :* We proceed as in the proof of Lemma 3.2 of Segers et al. (2017). Fix  $\eta > 0$  and let us show that, with probability one,  $J_n$  can be made smaller than  $\eta$  provided  $n$  is large enough. Let  $|\cdot|_\infty$  denote the maximum norm on  $\mathbb{R}^d$ . For any  $\varepsilon > 0$ , we have that

$$\begin{aligned} J_n &= \sup_{\mathbf{u} \in [0,1]^d} \left| \int_{[0,1]^d} \{C(\mathbf{w}) - C(\mathbf{u})\} d\nu_{\mathbf{u}}^{\mathcal{X}_{1:n}}(\mathbf{w}) \right| \\ &\leq \sup_{\mathbf{u} \in [0,1]^d} \left| \int_{\{\mathbf{w} \in [0,1]^d : |\mathbf{u} - \mathbf{w}|_\infty \leq \varepsilon\}} \{C(\mathbf{w}) - C(\mathbf{u})\} d\nu_{\mathbf{u}}^{\mathcal{X}_{1:n}}(\mathbf{w}) \right| \\ &\quad + \sup_{\mathbf{u} \in [0,1]^d} \left| \int_{\{\mathbf{w} \in [0,1]^d : |\mathbf{u} - \mathbf{w}|_\infty > \varepsilon\}} \{C(\mathbf{w}) - C(\mathbf{u})\} d\nu_{\mathbf{u}}^{\mathcal{X}_{1:n}}(\mathbf{w}) \right| \leq J'_n + J''_n, \end{aligned}$$

where

$$\begin{aligned} J'_n &= \sup_{\substack{(\mathbf{u}, \mathbf{w}) \in [0,1]^{2d} \\ |\mathbf{u} - \mathbf{w}|_\infty \leq \varepsilon}} |C(\mathbf{w}) - C(\mathbf{u})|, \\ J''_n &= \sup_{\mathbf{u} \in [0,1]^d} \nu_{\mathbf{u}}^{\mathcal{X}_{1:n}}(\{\mathbf{w} \in [0,1]^d : |\mathbf{u} - \mathbf{w}|_\infty > \varepsilon\}). \end{aligned}$$

Let  $\varepsilon = \eta/(2d)$ . Then, from the Lipschitz continuity of  $C$ ,  $J'_n \leq \eta/2$ . As far as  $J''_n$  is concerned,

conditionally on  $\mathbf{X}_1, \mathbf{X}_2, \dots$ , for almost any sequence  $\mathbf{X}_1, \mathbf{X}_2, \dots$ , using Chebyshev's inequality and Condition 4.5.8, we have that

$$\begin{aligned} J_n'' &= \sup_{\mathbf{u} \in [0,1]^d} \mathbb{P} \{ |\mathbf{W}_{\mathbf{u}}^{\mathbf{X}_{1:n}} - \mathbf{u}|_{\infty} > \varepsilon \mid \mathbf{X}_{1:n} \} \\ &= \sup_{\mathbf{u} \in [0,1]^d} \mathbb{P} \left[ \bigcup_{j=1}^d \{ |W_{j,u_j}^{\mathbf{X}_{1:n}} - u_j| > \varepsilon \} \mid \mathbf{X}_{1:n} \right] \\ &\leq \sup_{\mathbf{u} \in [0,1]^d} \sum_{j=1}^d \mathbb{P} \{ |W_{j,u_j}^{\mathbf{X}_{1:n}} - u_j| > \varepsilon \mid \mathbf{X}_{1:n} \} \\ &\leq \sup_{\mathbf{u} \in [0,1]^d} \sum_{j=1}^d \frac{\text{Var}(W_{j,u_j}^{\mathbf{X}_{1:n}} \mid \mathbf{X}_{1:n})}{\varepsilon^2} \leq \frac{da_n}{\varepsilon^2}. \end{aligned}$$

which implies that, for  $n$  sufficiently large,  $J_n'' \leq \eta/2$ . The latter holds conditionally on  $\mathbf{X}_1, \mathbf{X}_2, \dots$  for almost any sequence  $\mathbf{X}_1, \mathbf{X}_2, \dots$ , which completes the proof.  $\blacksquare$

Next, we recall the mode of convergence classically used to state asymptotic validity results of resampling techniques when dealing with empirical processes; see, e.g., [van der Vaart and Wellner \(2000, Chapter 2.9\)](#) or [Kosorok \(2008, Section 2.2.3\)](#). Let

$$BL_1 = \{h : \ell^\infty([0, 1]^d) \rightarrow [-1, 1] \text{ such that,}$$

$$\text{for all } x, y \in \ell^\infty([0, 1]^d), |h(x) - h(y)| \leq \sup_{\mathbf{u} \in [0,1]^d} |x(\mathbf{u}) - y(\mathbf{u})|\}.$$

Let  $\mathbb{X}_n = \mathbb{X}_n(\mathbf{X}_{1:n}, \mathcal{W}_n)$  be a sequence of bootstrapped empirical processes in  $\ell^\infty([0, 1]^d)$  depending on an additional source of randomness  $\mathcal{W}_n$  (often called the ‘‘bootstrap weights’’). For the smooth bootstraps under consideration,  $\mathcal{W}_n$  is independent of the data  $\mathbf{X}_{1:n}$  and corresponds to  $n$  independent copies of the independent random variables  $I$  and  $\mathbf{U}^\#$  necessary to carry out Algorithm 4.3.2 (see also (4.18))  $n$  times independently. The notation  $\mathbb{X}_n \xrightarrow[\mathcal{W}]{\mathbb{P}} \mathbb{X}$  then means that

- $\sup_{h \in BL_1} |\mathbb{E}_{\mathcal{W}}\{h(\mathbb{X}_n)\} - \mathbb{E}\{h(\mathbb{X})\}| \rightarrow 0$  in outer probability,
- $\mathbb{E}_{\mathcal{W}}\{h(\mathbb{X}_n)^*\} - \mathbb{E}_{\mathcal{W}}\{h(\mathbb{X}_n)_*\} \xrightarrow{P} 0$  for all  $h \in BL_1$ ,

where  $\mathbb{E}_{\mathcal{W}}$  denotes an expectation with respect to the bootstrap weights  $\mathcal{W}_n$  only and  $h(\mathbb{X}_n)^*$  and  $h(\mathbb{X}_n)_*$  denote the minimal measurable majorant and maximal measurable minorant with respect to  $(\mathbf{X}_{1:n}, \mathcal{W}_n)$ .

The next lemma is very closely related to Proposition 3.3 of [Kiriliouk et al. \(2021\)](#).

**Lemma 4.7.6.** *Assume that the random vectors in  $\mathbf{X}_{1:n}$  are i.i.d., and that Conditions 4.2.4, 4.2.5, 4.2.6, 4.2.8 and 4.2.9 hold. Then,*

$$\sqrt{n}(C_{1:n}^\# - C_{1:n}^\nu) \xrightarrow[\mathcal{W}]{\mathbb{P}} \mathbb{C}_C(0, 1, \cdot), \tag{4.65}$$

where  $\mathbb{C}_C$  is defined in (4.14).

*Proof.* Let  $G_{1:n}^\#$  be the empirical d.f. of  $\mathbf{V}_{1:n}^\#$ . Using Lemma 4.7.5 and proceeding as in Step 1 of the proof of Proposition 3.3 of Kiriliouk et al. (2021), one obtains that

$$\sqrt{n}(G_{1:n}^\# - C_{1:n}^\nu) \overset{\mathbb{P}}{\rightsquigarrow} \mathbb{B}_C(0, 1, \cdot), \quad (4.66)$$

where  $\mathbb{B}_C$  is defined in (4.12). Then, proceeding as in Step 2 of the proof of Proposition 3.3 of Kiriliouk et al. (2021), that is, combining (4.66) with the Hadamard differentiability of the map  $\Phi$  in (4.63) established in Theorem 2.4 of Bücher and Volgushev (2013), the functional delta method for the bootstrap “in probability” (van der Vaart and Wellner, 2000, Theorem 3.9.11) and the fact  $\Phi(C_{1:n}^\nu) = C_{1:n}^\nu$  (since  $C_{1:n}^\nu$  has standard uniform margins under Condition 4.2.4 in the considered i.i.d. setting), one obtains

$$\sqrt{n}(\Phi(G_{1:n}^\#) - C_{1:n}^\nu) \overset{\mathbb{P}}{\rightsquigarrow} \mathbb{C}_C(0, 1, \cdot). \quad (4.67)$$

The desired result finally follows from (4.67) and the well-known fact that

$$\sup_{\mathbf{u} \in [0,1]^d} |C_{1:n}^\#(\mathbf{u}) - \Phi(G_{1:n}^\#)(\mathbf{u})| \leq \frac{d}{n}$$

since the components samples of  $\mathbf{V}_{1:n}^\#$  contain no ties with probability one. ■

**Proof of Theorem 4.3.5.** Combining Lemma 4.7.6 with Lemma 3.1 of Bücher and Kojadinovic (2019), we obtain that (4.65) is equivalent to

$$(\mathbb{C}_n(0, 1, \cdot), \sqrt{n}(C_{1:n}^{\#, [1]} - C_{1:n}^\nu), \sqrt{n}(C_{1:n}^{\#, [2]} - C_{1:n}^\nu)) \rightsquigarrow (\mathbb{C}_C(0, 1, \cdot), \mathbb{C}_C^{[1]}(0, 1, \cdot), \mathbb{C}_C^{[2]}(0, 1, \cdot)) \quad (4.68)$$

in  $\{\ell^\infty([0, 1]^d)\}^3$ , where  $\mathbb{C}_n$  is defined in (4.9). From Theorem 4.2.10, we have that

$$\sup_{\mathbf{u} \in [0,1]^d} \sqrt{n}|C_{1:n}^\nu(\mathbf{u}) - C_{1:n}(\mathbf{u})| = \sup_{\mathbf{u} \in [0,1]^d} |\mathbb{C}_n^\nu(0, 1, \mathbf{u}) - \mathbb{C}_n(0, 1, \mathbf{u})| = o_{\mathbb{P}}(1), \quad (4.69)$$

where  $\mathbb{C}_n^\nu$  is defined in (4.10). The first joint weak convergence in Theorem 4.3.5 then follows from (4.68) and (4.69).

Fix  $j \in \{1, 2\}$ . Since (4.69) holds, to establish the second joint weak convergence from the first, it suffices to show that

$$\sup_{\mathbf{u} \in [0,1]^d} |\sqrt{n}\{C_{1:n}^{\#, \nu, [j]}(\mathbf{u}) - C_{1:n}^\nu(\mathbf{u})\} - \sqrt{n}\{C_{1:n}^{\#, [j]}(\mathbf{u}) - C_{1:n}(\mathbf{u})\}| = o_{\mathbb{P}}(1). \quad (4.70)$$



The supremum on the left hand-side of (4.70) is smaller than  $I_n + J_n$ , where

$$I_n = \sup_{\mathbf{u} \in [0,1]^d} \left| \int_{[0,1]^d} \sqrt{n} \{C_{1:n}^{\#, [j]}(\mathbf{w}) - C_{1:n}(\mathbf{w})\} d\nu_{\mathbf{u}}^{\#, [j]}(\mathbf{w}) - \sqrt{n} \{C_{1:n}^{\#, [j]}(\mathbf{u}) - C_{1:n}(\mathbf{u})\} \right|,$$

$$J_n = \sqrt{n} \sup_{\mathbf{u} \in [0,1]^d} \left| \int_{[0,1]^d} C_{1:n}(\mathbf{w}) d\nu_{\mathbf{u}}^{\#, [j]}(\mathbf{w}) - \int_{[0,1]^d} C_{1:n}(\mathbf{w}) d\nu_{\mathbf{u}}^{\mathbf{x}_{1:n}}(\mathbf{w}) \right|.$$

Since, according to the first claim of Theorem 4.3.5,  $\sqrt{n}(C_{1:n}^{\#, [j]} - C_{1:n}) \rightsquigarrow \mathbb{C}_C(0, 1, \cdot)$  in  $\ell^\infty([0, 1]^d)$  and  $\mathbb{C}_C(0, 1, \cdot)$  has continuous trajectories almost surely, it can be verified by proceeding as in the proof of Lemma 4.7.1 that  $I_n = o_{\mathbb{P}}(1)$ . For the term  $J_n$ , we have that  $J_n \leq K_n + L_n$ , where

$$K_n = \sqrt{n} \sup_{\mathbf{u} \in [0,1]^d} \left| \int_{[0,1]^d} \{C_{1:n}(\mathbf{w}) - C(\mathbf{w})\} d\nu_{\mathbf{u}}^{\#, [j]}(\mathbf{w}) - \int_{[0,1]^d} \{C_{1:n}(\mathbf{w}) - C(\mathbf{w})\} d\nu_{\mathbf{u}}^{\mathbf{x}_{1:n}}(\mathbf{w}) \right|,$$

$$L_n = \sqrt{n} \sup_{\mathbf{u} \in [0,1]^d} \left| \int_{[0,1]^d} C(\mathbf{w}) d\nu_{\mathbf{u}}^{\#, [j]}(\mathbf{w}) - \int_{[0,1]^d} C(\mathbf{w}) d\nu_{\mathbf{u}}^{\mathbf{x}_{1:n}}(\mathbf{w}) \right|.$$

The term  $K_n$  is smaller than  $K'_n + K''_n$ , where

$$K'_n = \sup_{\mathbf{u} \in [0,1]^d} \left| \int_{[0,1]^d} \mathbb{C}_n(0, 1, \mathbf{w}) d\nu_{\mathbf{u}}^{\#, [j]}(\mathbf{w}) - \mathbb{C}_n(0, 1, \mathbf{u}) \right|,$$

$$K''_n = \sup_{\mathbf{u} \in [0,1]^d} \left| \mathbb{C}_n(0, 1, \mathbf{u}) - \int_{[0,1]^d} \mathbb{C}_n(0, 1, \mathbf{w}) d\nu_{\mathbf{u}}^{\mathbf{x}_{1:n}}(\mathbf{w}) \right|.$$

From (4.57) in Lemma 4.7.1,  $K''_n = o_{\mathbb{P}}(1)$  and, proceeding again as in the proof of the latter lemma, it can be verified that  $K'_n = o_{\mathbb{P}}(1)$ . The term  $L_n$  is smaller than  $L'_n + L''_n$ , where

$$L'_n = \sqrt{n} \sup_{\mathbf{u} \in [0,1]^d} \left| \int_{[0,1]^d} C(\mathbf{w}) d\nu_{\mathbf{u}}^{\#, [j]}(\mathbf{w}) - C(\mathbf{u}) \right|,$$

$$L''_n = \sqrt{n} \sup_{\mathbf{u} \in [0,1]^d} \left| C(\mathbf{u}) - \int_{[0,1]^d} C(\mathbf{w}) d\nu_{\mathbf{u}}^{\mathbf{x}_{1:n}}(\mathbf{w}) \right|.$$

The term  $L''_n$  converges almost surely to zero as a consequence of (4.59) in Lemma 4.7.2. The proof of the latter result can be adapted to verify that the term  $L'_n$  also converges almost surely to zero. Hence, (4.70) holds, which completes the proof.  $\blacksquare$

#### 4.7.4 Proof of Theorem 4.4.3

The proof of Theorem 4.4.3 is based on the following lemma which we prove first.

**Lemma 4.7.7.** *Under Conditions 4.4.2 and 4.5.8 (the latter is implied by Condition 4.2.9), for any  $b \in \mathbb{N}$ ,*

$$\sup_{(s,t,\mathbf{u}) \in \Lambda \times [0,1]^d} \left| \hat{\mathbb{B}}_n^{[b],\nu}(s,t,\mathbf{u}) - \hat{\mathbb{B}}_n^{[b]}(s,t,\mathbf{u}) \right| = o_{\mathbb{P}}(1), \quad (4.71)$$

$$\sup_{(s,t,\mathbf{u}) \in \Lambda \times [0,1]^d} \left| \check{\mathbb{B}}_n^{[b],\nu}(s,t,\mathbf{u}) - \check{\mathbb{B}}_n^{[b]}(s,t,\mathbf{u}) \right| = o_{\mathbb{P}}(1). \quad (4.72)$$

*Proof.* Fix  $b \in \mathbb{N}$ . We first prove (4.71). Starting from (4.26), we have that

$$\begin{aligned} & \sup_{(s,t,\mathbf{u}) \in \Lambda \times [0,1]^d} \left| \hat{\mathbb{B}}_n^{[b],\nu}(s,t,\mathbf{u}) - \hat{\mathbb{B}}_n^{[b]}(s,t,\mathbf{u}) \right| \\ &= \sup_{(s,t,\mathbf{u}) \in \Lambda \times [0,1]^d} \left| \int_{[0,1]^d} \hat{\mathbb{B}}_n^{[b]}(s,t,\mathbf{w}) d\nu_{\mathbf{u}}^{\mathbf{x}^{1:n}}(\mathbf{w}) - \hat{\mathbb{B}}_n^{[b]}(s,t,\mathbf{u}) \right| = o_{\mathbb{P}}(1), \end{aligned}$$

where the last equality follows from (4.57) in Lemma 4.7.1 since  $\hat{\mathbb{B}}_n^{[b]} \rightsquigarrow \mathbb{B}_C$  in  $\ell^\infty(\Lambda \times [0,1]^d)$ , where  $\mathbb{B}_C$  is defined in (4.12) and has continuous trajectories almost surely under Condition 4.4.2. Under Condition 4.4.2 (ii), the latter is a consequence of Lemmas D.1 and D.2 in Bücher and Kojadinovic (2016) as well as Theorem 2.1 in Bücher and Kojadinovic (2016). Under Condition 4.4.2 (i), one can rely on Theorem 1 of Holmes et al. (2013) instead.

The proof of (4.72) is similar. Starting from (4.27), we have that

$$\begin{aligned} & \sup_{(s,t,\mathbf{u}) \in \Lambda \times [0,1]^d} \left| \check{\mathbb{B}}_n^{[b],\nu}(s,t,\mathbf{u}) - \check{\mathbb{B}}_n^{[b]}(s,t,\mathbf{u}) \right| \\ &= \sup_{(s,t,\mathbf{u}) \in \Lambda \times [0,1]^d} \left| \int_{[0,1]^d} \check{\mathbb{B}}_n^{[b]}(s,t,\mathbf{w}) d\nu_{\mathbf{u}}^{\mathbf{x}^{[ns]+1:[nt]}}(\mathbf{w}) - \check{\mathbb{B}}_n^{[b]}(s,t,\mathbf{u}) \right| = o_{\mathbb{P}}(1), \end{aligned}$$

where the last equality follows from (4.56) in Lemma 4.7.1 since  $\check{\mathbb{B}}_n^{[b]} \rightsquigarrow \mathbb{B}_C$  in  $\ell^\infty(\Lambda \times [0,1]^d)$ . Under Condition 4.4.2 (ii), the latter is a consequence of (B.3) in the proof of Proposition 4.3 in Bücher et al. (2014) and Theorem 2.1 in Bücher and Kojadinovic (2016). Under Condition 4.4.2 (i), one can rely again on Theorem 1 of Holmes et al. (2013) instead. ■

**Proof of Theorem 4.4.3.** Fix  $b \in \mathbb{N}$ . We only prove (4.31), the proof of (4.30) being simpler. Starting from (4.25) and (4.29), we have that

$$\begin{aligned} & \sup_{\substack{(s,t) \in \Lambda \\ \mathbf{u} \in [0,1]^d}} \left| \check{\mathbb{C}}_n^{[b]}(s,t,\mathbf{u}) - \check{\mathbb{C}}_n^{[b],\nu}(s,t,\mathbf{u}) \right| \leq \sup_{\substack{(s,t) \in \Lambda \\ \mathbf{u} \in [0,1]^d}} \left| \check{\mathbb{B}}_n^{[b]}(s,t,\mathbf{u}) - \check{\mathbb{B}}_n^{[b],\nu}(s,t,\mathbf{u}) \right| \\ &+ \sum_{j=1}^d \sup_{\substack{(s,t) \in \Lambda \\ \mathbf{u} \in [0,1]^d}} \left| \dot{C}_{j,[ns]+1:[nt]}(\mathbf{u}) \right| \sup_{\substack{(s,t) \in \Lambda \\ \mathbf{u} \in [0,1]^d}} \left| \check{\mathbb{B}}_n^{[b]}(s,t,\mathbf{u}^{(j)}) - \check{\mathbb{B}}_n^{[b],\nu}(s,t,\mathbf{u}^{(j)}) \right|. \end{aligned}$$

The terms on the right-hand side of the previous display converge to zero in probability as a consequence of (4.72) in Lemma 4.7.7 and the fact that  $\sup_{(s,t,\mathbf{u}) \in \Lambda \times [0,1]^d} \left| \dot{C}_{j, [ns]+1: [nt]}(\mathbf{u}) \right| \leq \zeta$  from Condition 4.4.1.

The last two claims of the theorem are an immediate consequence of (4.30), (4.31) and straightforward extensions of Propositions 4.2 and 4.3 in Bücher et al. (2014) for non-smooth multiplier replicates based on arbitrary partial derivative estimators satisfying Condition 4.4.1. ■

#### 4.7.5 Proofs of Propositions 4.5.3, 4.5.5, 4.5.9 and 4.5.10

**Lemma 4.7.8.** *Let  $f$  be any function from  $[0, 1]^d$  to  $[0, 1]$ , let  $m \in \mathbb{N}$ ,  $m \geq 2$  and recall that, for any  $\mathbf{u} \in [0, 1]^d$ ,  $\mu_{m,\mathbf{u}}$  is the law of the random vector  $(S_{m,1,u_1}/m, \dots, S_{m,d,u_d}/m)$ , where  $S_{m,1,u_1}, \dots, S_{m,d,u_d}$  are independent, and for each  $k \in \{1, \dots, d\}$ ,  $S_{m,k,u_k}$  is Binomial( $m, u_k$ ). Moreover, recall that, for any  $j \in \{1, \dots, d\}$ ,  $\tilde{\mu}_{j,m,\mathbf{u}}$  is the law of the random vector  $(\tilde{S}_{m,1,u_1}/m, \dots, \tilde{S}_{m,d,u_d}/m)$  whose components are independent and, for  $i \in \{1, \dots, d\} \setminus \{j\}$ ,  $\tilde{S}_{m,i,u_i}$  is Binomial( $m, u_i$ ), whereas  $\tilde{S}_{m,j,u_j}$  is Binomial( $m - 1, u_j$ ). Then, for any  $j \in \{1, \dots, d\}$  and  $\mathbf{u} \in [0, 1]^d$  such that  $u_j \in (0, 1)$ ,*

$$\partial_{u_j} \left\{ \int_{[0,1]^d} f(\mathbf{w}) d\mu_{m,\mathbf{u}}(\mathbf{w}) \right\} = m \int_{[0,1]^d} \{f(\mathbf{w} + \mathbf{e}_j/m) - f(\mathbf{w})\} d\tilde{\mu}_{j,m,\mathbf{u}}(\mathbf{w}).$$

*Proof.* Fix  $m \geq 2$  and, without loss of generality, fix  $j = 1$ . Also, for any  $u \in [0, 1]$ , let  $b_{m,u}(s) = \binom{m}{s} u^s (1-u)^{m-s}$ ,  $s \in \{0, \dots, m\}$ . Then, for all  $\mathbf{u} \in [0, 1]^d$  such that  $u_1 \in (0, 1)$ ,

$$\begin{aligned}
 & \partial_{u_1} \left\{ \int_{[0,1]^d} f(\mathbf{w}) d\mu_{m,\mathbf{u}}(\mathbf{w}) \right\} = \partial_{u_1} \left\{ \sum_{s_1=0}^m \cdots \sum_{s_d=0}^m f\left(\frac{s_1}{m}, \dots, \frac{s_d}{m}\right) \prod_{j=1}^d b_{m,u_j}(s_j) \right\} \\
 &= \sum_{s_1=0}^m \cdots \sum_{s_d=0}^m f\left(\frac{s_1}{m}, \dots, \frac{s_d}{m}\right) \partial_{u_1} b_{m,u_1}(s_1) \prod_{j=2}^d b_{m,u_j}(s_j) \\
 &= \sum_{s_1=0}^m \cdots \sum_{s_d=0}^m f\left(\frac{s_1}{m}, \dots, \frac{s_d}{m}\right) \\
 & \quad \times \binom{m}{s_1} \left\{ s_1 u_1^{s_1-1} (1-u_1)^{m-s_1} - (m-s_1) u_1^{s_1} (1-u_1)^{m-s_1-1} \right\} \prod_{j=2}^d b_{m,u_j}(s_j) \\
 &= m \sum_{s_1=1}^m \cdots \sum_{s_d=0}^m f\left(\frac{s_1}{m}, \dots, \frac{s_d}{m}\right) \frac{(m-1)!}{(s_1-1)!(m-s_1)!} u_1^{s_1-1} (1-u_1)^{m-s_1} \prod_{j=2}^d b_{m,u_j}(s_j) \\
 & \quad - m \sum_{s_1=0}^{m-1} \cdots \sum_{s_d=0}^m f\left(\frac{s_1}{m}, \dots, \frac{s_d}{m}\right) \frac{(m-1)!}{s_1!(m-s_1-1)!} u_1^{s_1} (1-u_1)^{m-s_1-1} \prod_{j=2}^d b_{m,u_j}(s_j) \\
 &= m \sum_{s_1=0}^{m-1} \cdots \sum_{s_d=0}^m f\left(\frac{s_1+1}{m}, \dots, \frac{s_d}{m}\right) \frac{(m-1)!}{s_1!(m-s_1-1)!} u_1^{s_1} (1-u_1)^{m-s_1-1} \prod_{j=2}^d b_{m,u_j}(s_j) \\
 & \quad - m \sum_{s_1=0}^{m-1} \cdots \sum_{s_d=0}^m f\left(\frac{s_1}{m}, \dots, \frac{s_d}{m}\right) \frac{(m-1)!}{s_1!(m-s_1-1)!} u_1^{s_1} (1-u_1)^{m-s_1-1} \prod_{j=2}^d b_{m,u_j}(s_j) \\
 &= m \sum_{s_1=0}^{m-1} \cdots \sum_{s_d=0}^m \left\{ f\left(\frac{s_1+1}{m}, \dots, \frac{s_d}{m}\right) - f\left(\frac{s_1}{m}, \dots, \frac{s_d}{m}\right) \right\} b_{m-1,u_1}(s_1) \prod_{j=2}^d b_{m,u_j}(s_j) \\
 &= m \int_{[0,1]^d} \{f(\mathbf{w} + \mathbf{e}_1/m) - f(\mathbf{w})\} d\tilde{\mu}_{1,m,\mathbf{u}}(\mathbf{w}).
 \end{aligned}$$

■

**Proof of Proposition 4.5.3.** Fix  $j \in \{1, \dots, d\}$ ,  $\mathbf{u} \in [0, 1]^d$  and  $m \geq 2$ , and recall the definition of the measure  $\tilde{\mu}_{j,m,\mathbf{u}}$  given in Lemma 4.7.8. From (4.50), we have that

$$\dot{C}_{j,k:l,m}^{\text{Bern}}(\mathbf{u}) = m \int_{[0,1]^d} C_{k:l}(\mathbf{w} + \mathbf{e}_j/m) d\tilde{\mu}_{j,m,\mathbf{u}}(\mathbf{w}) - m \int_{[0,1]^d} C_{k:l}(\mathbf{w}) d\tilde{\mu}_{j,m,\mathbf{u}}(\mathbf{w}). \quad (4.73)$$

Let  $\tilde{\mathbf{S}} = (\tilde{S}_{m,1,u_1}, \dots, \tilde{S}_{m,d,u_d})$  so that  $\tilde{\mathbf{S}}/m$  is a random vector with law  $\tilde{\mu}_{j,m,\mathbf{u}}$ . Then, combined with the definition of  $C_{k:l}$  in (4.2), the first integral on the right-hand side of (4.73) can be

rewritten as

$$\begin{aligned}
 & \int_{[0,1]^d} \frac{1}{l-k+1} \sum_{i=k}^l \mathbf{1} \left\{ \mathbf{R}_i^{k:l} / (l-k+1) \leq \mathbf{w} + \mathbf{e}_j / m \right\} d\tilde{\mu}_{j,m,\mathbf{u}}(\mathbf{w}) \\
 &= \frac{1}{l-k+1} \sum_{i=k}^l \int_{[0,1]^d} \mathbf{1} \left\{ \mathbf{R}_i^{k:l} / (l-k+1) - \mathbf{e}_j / m \leq \mathbf{w} \right\} d\tilde{\mu}_{j,m,\mathbf{u}}(\mathbf{w}) \\
 &= \frac{1}{l-k+1} \sum_{i=k}^l \mathbb{P} \left\{ \tilde{\mathbf{S}} \geq m \mathbf{R}_i^{k:l} / (l-k+1) - \mathbf{e}_j \mid \boldsymbol{\mathcal{X}}_{k:l} \right\} \\
 &= \frac{1}{l-k+1} \sum_{i=k}^l \mathbb{P} \left\{ \tilde{S}_{m,j,u_j} \geq m R_{ij}^{k:l} / (l-k+1) - 1 \mid \boldsymbol{\mathcal{X}}_{k:l} \right\} \\
 &\quad \times \prod_{\substack{t=1 \\ t \neq j}}^d \mathbb{P} \left\{ \tilde{S}_{m,t,u_t} \geq m R_{it}^{k:l} / (l-k+1) \mid \boldsymbol{\mathcal{X}}_{k:l} \right\} \\
 &= \frac{1}{l-k+1} \sum_{i=k}^l \mathbb{P} \left\{ \tilde{S}_{m,j,u_j} > \lceil m R_{ij}^{k:l} / (l-k+1) - 1 \rceil - 1 \mid \boldsymbol{\mathcal{X}}_{k:l} \right\} \\
 &\quad \times \prod_{\substack{t=1 \\ t \neq j}}^d \mathbb{P} \left\{ \tilde{S}_{m,t,u_t} > \lceil m R_{it}^{k:l} / (l-k+1) \rceil - 1 \mid \boldsymbol{\mathcal{X}}_{k:l} \right\} \\
 &= \frac{1}{l-k+1} \sum_{i=k}^l \bar{B}_{m-1,u_j} \left\{ \lceil m R_{ij}^{k:l} / (l-k+1) \rceil - 2 \right\} \\
 &\quad \times \prod_{\substack{t=1 \\ t \neq j}}^d \bar{B}_{m,u_t} \left\{ \lceil m R_{it}^{k:l} / (l-k+1) \rceil - 1 \right\},
 \end{aligned}$$

where we have used the fact that, for any  $t \in \{1, \dots, d\}$  and  $x \in \mathbb{R}$ ,  $\mathbb{P}(\tilde{S}_{m,t,u_t} \geq x) = \mathbb{P}(\tilde{S}_{m,t,u_t} > \lceil x \rceil - 1)$  and  $\lceil x - 1 \rceil = \lceil x \rceil - 1$ . Similarly, for the second integral on the right-hand side of (4.73),

we have

$$\begin{aligned}
 & \int_{[0,1]^d} C_{k:l}(\mathbf{w}) d\tilde{\mu}_{j,m,\mathbf{u}}(\mathbf{w}) \\
 &= \frac{1}{l-k+1} \int_{[0,1]^d} \sum_{i=k}^l \mathbf{1} \left\{ \mathbf{R}_i^{k:l} / (l-k+1) \leq \mathbf{w} \right\} d\tilde{\mu}_{j,m,\mathbf{u}}(\mathbf{w}) \\
 &= \frac{1}{l-k+1} \sum_{i=k}^l \mathbb{P} \left\{ \tilde{\mathcal{S}} \geq m\mathbf{R}_i^{k:l} / (l-k+1) \mid \mathcal{X}_{k:l} \right\} \\
 &= \frac{1}{l-k+1} \sum_{i=k}^l \prod_{t=1}^d \mathbb{P} \left\{ \tilde{S}_{m,t,u_t} \geq mR_{it}^{k:l} / (l-k+1) \mid \mathcal{X}_{k:l} \right\} \\
 &= \frac{1}{l-k+1} \sum_{i=k}^l \bar{B}_{m-1,u_j} \left\{ \lceil mR_{ij}^{k:l} / (l-k+1) \rceil - 1 \right\} \\
 &\quad \times \prod_{\substack{t=1 \\ t \neq j}}^d \bar{B}_{m,u_t} \left\{ \lceil mR_{it}^{k:l} / (l-k+1) \rceil - 1 \right\}.
 \end{aligned}$$

The desired result finally follows by noticing that

$$\begin{aligned}
 & \bar{B}_{m-1,u_j} \left\{ \lceil mR_{ij}^{k:l} / (l-k+1) \rceil - 2 \right\} - \bar{B}_{m-1,u_j} \left\{ \lceil mR_{ij}^{k:l} / (l-k+1) \rceil - 1 \right\} \\
 &= B_{m-1,u_j} \left\{ \lceil mR_{ij}^{k:l} / (l-k+1) \rceil - 1 \right\} - B_{m-1,u_j} \left\{ \lceil mR_{ij}^{k:l} / (l-k+1) \rceil - 2 \right\} \\
 &= b_{m-1,u_j} \left\{ \lceil mR_{ij}^{k:l} / (l-k+1) \rceil - 1 \right\}.
 \end{aligned}$$

■

**Lemma 4.7.9.** *Under Condition 4.5.4, for any  $j \in \{1, \dots, d\}$ ,  $\delta \in (0, 1)$  and  $\varepsilon \in (0, 1/2)$ , with probability 1,*

$$\sup_{\substack{(s,t) \in \Lambda \\ t-s \geq \delta}} \sup_{\substack{\mathbf{u} \in [0,1]^d \\ u_j \in [\varepsilon, 1-\varepsilon]}} \left| \dot{C}_{j,[ns]+1:[nt]}^{\nu,\nabla}(\mathbf{u}) - \dot{C}_{j,[ns]+1:[nt]}^{\nu,\Delta}(\mathbf{u}) \right| = o(1).$$

*Proof.* Fix  $j \in \{1, \dots, d\}$ ,  $\delta \in (0, 1)$  as well as  $\varepsilon \in (0, 1/2)$  and assume that  $n$  is large enough so that, for any  $(s, t) \in \Lambda$  such that  $t - s > \delta$ ,  $L_2 b_{[nt]-[ns]}$  and  $L_2 b'_{[nt]-[ns]}$  are smaller than  $\varepsilon$ . Then, using the fact that  $C_{k:l}^\nu$  in (4.3) is between 0 and 1, we obtain that, with probability 1,

$$\begin{aligned}
 & \sup_{\substack{(s,t) \in \Lambda \\ t-s \geq \delta}} \sup_{\substack{\mathbf{u} \in [0,1]^d \\ u_j \in [\varepsilon, 1-\varepsilon]}} \left| \dot{C}_{j,[ns]+1:[nt]}^{\nu,\nabla}(\mathbf{u}) - \dot{C}_{j,[ns]+1:[nt]}^{\nu,\Delta}(\mathbf{u}) \right| \\
 & \leq \sup_{\substack{(s,t) \in \Lambda \\ t-s \geq \delta}} \sup_{\substack{\mathbf{u} \in [0,1]^d \\ u_j \in [\varepsilon, 1-\varepsilon]}} \left| \frac{1}{h+h'} - \frac{1}{(u_j+h) \wedge 1 - (u_j-h') \vee 0} \right| = 0.
 \end{aligned}$$

■

**Proof of Proposition 4.5.5.** Fix  $j \in \{1, \dots, d\}$  and let us first prove (4.52) by proceeding along the lines of the proof of (B.4) in Bücher et al. (2014). From (4.10), notice that, for any  $(s, t, \mathbf{u}) \in \Lambda \times [0, 1]^d$  such that  $\lfloor ns \rfloor < \lfloor nt \rfloor$ ,

$$C_{\lfloor ns \rfloor + 1: \lfloor nt \rfloor}^\nu(\mathbf{u}) = C(\mathbf{u}) + \frac{1}{\sqrt{n}\lambda_n(s, t)} \mathbb{C}_n^\nu(s, t, \mathbf{u}).$$

Fix  $\delta \in (0, 1)$  and notice that, by Condition 4.5.4,

$$d_n = \sup_{\substack{(s, t) \in \Lambda \\ t-s \geq \delta}} (b_{\lfloor nt \rfloor - \lfloor ns \rfloor} + b'_{\lfloor nt \rfloor - \lfloor ns \rfloor}) \leq \sup_{k \geq \lfloor n\delta \rfloor - 1} (b_k + b'_k) \rightarrow 0. \quad (4.74)$$

Next, fix  $\varepsilon \in (0, 1/2)$  and assume that  $n$  is large enough so that, for any  $t - s > \delta$ ,  $L_2 b_{\lfloor nt \rfloor - \lfloor ns \rfloor}$  and  $L_2 b'_{\lfloor nt \rfloor - \lfloor ns \rfloor}$  are smaller than  $\varepsilon/2$ . Then, for any  $t - s > \delta$  and  $\mathbf{u} \in [0, 1]^d$  such that  $u_j \in [\varepsilon, 1 - \varepsilon]$ ,

$$\begin{aligned} \dot{C}_{j, \lfloor ns \rfloor + 1: \lfloor nt \rfloor}^{\nu, \Delta}(\mathbf{u}) &= \frac{1}{h + h'} \{C(\mathbf{u} + h\mathbf{e}_j) - C(\mathbf{u} - h'\mathbf{e}_j)\} \\ &\quad + \frac{1}{(h + h')\sqrt{n}\lambda_n(s, t)} \{\mathbb{C}_n^\nu(s, t, \mathbf{u} + h\mathbf{e}_j) - \mathbb{C}_n^\nu(s, t, \mathbf{u} - h'\mathbf{e}_j)\}. \end{aligned} \quad (4.75)$$

Since, by Condition 4.2.8,  $\dot{C}_j$  exists (and is continuous) on the set  $\{\mathbf{u} \in [0, 1]^d : u_j \in [\varepsilon/2, 1 - \varepsilon/2]\}$ , from the mean value theorem, for any  $t - s > \delta$  and  $\mathbf{u} \in [0, 1]^d$  such that  $u_j \in [\varepsilon, 1 - \varepsilon]$ ,

$$\frac{1}{h + h'} \{C(\mathbf{u} + h\mathbf{e}_j) - C(\mathbf{u} - h'\mathbf{e}_j)\} = \dot{C}_j(\mathbf{u}_{n, s, t}^*),$$

where  $\mathbf{u}_{n, s, t}^*$  is between  $\mathbf{u} - h'\mathbf{e}_j$  and  $\mathbf{u} + h\mathbf{e}_j$  almost surely. Hence, with probability 1,

$$\begin{aligned} &\sup_{\substack{(s, t) \in \Lambda \\ t-s \geq \delta}} \sup_{\substack{\mathbf{u} \in [0, 1]^d \\ u_j \in [\varepsilon, 1-\varepsilon]}} \left| \frac{1}{h + h'} \{C(\mathbf{u} + h\mathbf{e}_j) - C(\mathbf{u} - h'\mathbf{e}_j)\} - \dot{C}_j(\mathbf{u}) \right| \\ &= \sup_{\substack{(s, t) \in \Lambda \\ t-s \geq \delta}} \sup_{\substack{\mathbf{u} \in [0, 1]^d \\ u_j \in [\varepsilon, 1-\varepsilon]}} \left| \dot{C}_j(\mathbf{u}_{n, s, t}^*) - \dot{C}_j(\mathbf{u}) \right| \leq \sup_{\substack{(\mathbf{u}, \mathbf{v}) \in [0, 1]^{2d} \\ u_j, v_j \in [\varepsilon/2, 1-\varepsilon/2] \\ \|\mathbf{u} - \mathbf{v}\|_\infty \leq L_2 d_n}} \left| \dot{C}_j(\mathbf{u}) - \dot{C}_j(\mathbf{v}) \right| \rightarrow 0, \end{aligned} \quad (4.76)$$

where  $d_n$  is defined in (4.74). Furthermore, since, by Condition 4.2.9 and as a result of Theorem 4.2.10,  $\mathbb{C}_n^\nu$  is asymptotically uniformly equicontinuous in probability, we have that

$$\begin{aligned} &\sup_{\substack{(s, t) \in \Lambda \\ t-s \geq \delta}} \sup_{\substack{\mathbf{u} \in [0, 1]^d \\ u_j \in [\varepsilon, 1-\varepsilon]}} \left| \mathbb{C}_n^\nu(s, t, \mathbf{u} + h\mathbf{e}_j) - \mathbb{C}_n^\nu(s, t, \mathbf{u} - h'\mathbf{e}_j) \right| \\ &\leq \sup_{\substack{(s, t) \in \Lambda \\ t-s \geq \delta}} \sup_{\substack{(\mathbf{u}, \mathbf{v}) \in [0, 1]^{2d} \\ u_j, v_j \in [\varepsilon/2, 1-\varepsilon/2] \\ \|\mathbf{u} - \mathbf{v}\|_\infty \leq L_2 d_n}} \left| \mathbb{C}_n^\nu(s, t, \mathbf{u}) - \mathbb{C}_n^\nu(s, t, \mathbf{v}) \right| = o_{\mathbb{P}}(1). \end{aligned} \quad (4.77)$$

The fact that (4.52) holds is then an immediate consequence of (4.75), (4.76), (4.77) and the fact that, from Condition 4.5.4,

$$\begin{aligned} \sup_{\substack{(s,t) \in \Lambda \\ t-s \geq \delta}} \frac{1}{(h+h')\sqrt{n}\lambda_n(s,t)} &\leq \sup_{\substack{(s,t) \in \Lambda \\ t-s \geq \delta}} \frac{1}{L_1(b_{\lfloor nt \rfloor - \lfloor ns \rfloor} + b'_{\lfloor nt \rfloor - \lfloor ns \rfloor})\sqrt{n}\lambda_n(s,t)} \\ &\leq \sup_{\substack{(s,t) \in \Lambda \\ t-s \geq \delta}} \frac{1}{L_1(\lfloor nt \rfloor - \lfloor ns \rfloor)^{-1/2}\sqrt{n}\lambda_n(s,t)} \\ &= \sup_{\substack{(s,t) \in \Lambda \\ t-s \geq \delta}} \frac{1}{L_1\sqrt{\lambda_n(s,t)}} \leq \frac{1}{L_1\sqrt{\delta - 1/n}}. \end{aligned}$$

The claim for  $\dot{C}_{j,k;l}^{\nu,\nabla}$  (resp. for  $\dot{C}_{j,k;l}^{\nu,\Delta}$  and  $\dot{C}_{j,k;l}^{\nu,\nabla}$ ) follows from Lemma 4.7.9 (resp. the continuous mapping theorem).  $\blacksquare$

**Proof of Proposition 4.5.9.** Fix  $j \in \{1, \dots, d\}$ ,  $\delta \in (0, 1)$  and  $\varepsilon \in (0, 1/2)$ . We first prove (4.53). From (4.49) and the triangle inequality, we have that

$$\sup_{\substack{(s,t) \in \Lambda \\ t-s \geq \delta}} \sup_{\substack{\mathbf{u} \in [0,1]^d \\ u_j \in [\varepsilon, 1-\varepsilon]}} \left| \dot{C}_{j, \lfloor ns \rfloor + 1: \lfloor nt \rfloor}^{\Delta, \nu}(\mathbf{u}) - \dot{C}_j(\mathbf{u}) \right| \leq I_{j,n,\delta,\varepsilon} + J_{j,n,\delta,\varepsilon},$$

where

$$\begin{aligned} I_{j,n,\delta,\varepsilon} &= \sup_{\substack{(s,t) \in \Lambda \\ t-s \geq \delta}} \sup_{\substack{\mathbf{u} \in [0,1]^d \\ u_j \in [\varepsilon, 1-\varepsilon]}} \left| \int_{[0,1]^d} \left\{ \dot{C}_{j, \lfloor ns \rfloor + 1: \lfloor nt \rfloor}^{\Delta}(\mathbf{w}) - \dot{C}_j(\mathbf{w}) \right\} d\nu_{\mathbf{u}}^{\mathbf{x}_{\lfloor ns \rfloor + 1: \lfloor nt \rfloor}}(\mathbf{w}) \right|, \\ J_{j,n,\delta,\varepsilon} &= \sup_{\substack{(s,t) \in \Lambda \\ t-s \geq \delta}} \sup_{\substack{\mathbf{u} \in [0,1]^d \\ u_j \in [\varepsilon, 1-\varepsilon]}} \left| \int_{[0,1]^d} \dot{C}_j(\mathbf{w}) d\nu_{\mathbf{u}}^{\mathbf{x}_{\lfloor ns \rfloor + 1: \lfloor nt \rfloor}}(\mathbf{w}) - \dot{C}_j(\mathbf{u}) \right|, \end{aligned}$$

where  $\dot{C}_{j,k;l}^{\Delta}$  is defined in (4.41). We shall now show that both  $I_{j,n,\delta,\varepsilon} = o_{\mathbb{P}}(1)$  and  $J_{j,n,\delta,\varepsilon} = o_{\mathbb{P}}(1)$ .

*Term  $I_{j,n,\delta,\varepsilon}$ :* From the triangle inequality and the fact that  $0 \leq \dot{C}_{j,k;l}^{\Delta} \leq 1$  and  $0 \leq \dot{C}_j \leq 1$ , we have that  $I_{j,n,\delta,\varepsilon}$  is smaller than

$$\begin{aligned} &\sup_{\substack{(s,t) \in \Lambda \\ t-s \geq \delta}} \sup_{\substack{\mathbf{u} \in [0,1]^d \\ u_j \in [\varepsilon, 1-\varepsilon]}} \left| \int_{\{\mathbf{w} \in [0,1]^d: w_j \in [\varepsilon/2, 1-\varepsilon/2]\}} \left\{ \dot{C}_{j, \lfloor ns \rfloor + 1: \lfloor nt \rfloor}^{\Delta}(\mathbf{w}) - \dot{C}_j(\mathbf{w}) \right\} d\nu_{\mathbf{u}}^{\mathbf{x}_{\lfloor ns \rfloor + 1: \lfloor nt \rfloor}}(\mathbf{w}) \right| \\ &+ \sup_{\substack{(s,t) \in \Lambda \\ t-s \geq \delta}} \sup_{\substack{\mathbf{u} \in [0,1]^d \\ u_j \in [\varepsilon, 1-\varepsilon]}} \left| \int_{\{\mathbf{w} \in [0,1]^d: w_j < \varepsilon/2\}} \left\{ \dot{C}_{j, \lfloor ns \rfloor + 1: \lfloor nt \rfloor}^{\Delta}(\mathbf{w}) - \dot{C}_j(\mathbf{w}) \right\} d\nu_{\mathbf{u}}^{\mathbf{x}_{\lfloor ns \rfloor + 1: \lfloor nt \rfloor}}(\mathbf{w}) \right| \\ &+ \sup_{\substack{(s,t) \in \Lambda \\ t-s \geq \delta}} \sup_{\substack{\mathbf{u} \in [0,1]^d \\ u_j \in [\varepsilon, 1-\varepsilon]}} \left| \int_{\{\mathbf{w} \in [0,1]^d: w_j > 1-\varepsilon/2\}} \left\{ \dot{C}_{j, \lfloor ns \rfloor + 1: \lfloor nt \rfloor}^{\Delta}(\mathbf{w}) - \dot{C}_j(\mathbf{w}) \right\} d\nu_{\mathbf{u}}^{\mathbf{x}_{\lfloor ns \rfloor + 1: \lfloor nt \rfloor}}(\mathbf{w}) \right| \\ &\leq I'_{j,n,\delta,\varepsilon} + I''_{j,n,\delta,\varepsilon} + I'''_{j,n,\delta,\varepsilon}, \end{aligned}$$



where

$$\begin{aligned}
 I'_{j,n,\delta,\varepsilon} &= \sup_{\substack{(s,t) \in \Lambda \\ t-s \geq \delta}} \sup_{\substack{\mathbf{w} \in [0,1]^d \\ w_j \in [\varepsilon/2, 1-\varepsilon/2]}} \left| \dot{C}_{j, [ns]+1: [nt]}^\Delta(\mathbf{w}) - \dot{C}_j(\mathbf{w}) \right|, \\
 I''_{j,n,\delta,\varepsilon} &= \sup_{\substack{(s,t) \in \Lambda \\ t-s \geq \delta}} \sup_{\substack{\mathbf{u} \in [0,1]^d \\ u_j \in [\varepsilon, 1-\varepsilon]}} \nu_{\mathbf{u}}^{\mathbf{x}_{[ns]+1: [nt]}} \left\{ \mathbf{w} \in [0,1]^d : w_j < \varepsilon/2 \right\}, \\
 I'''_{j,n,\delta,\varepsilon} &= \sup_{\substack{(s,t) \in \Lambda \\ t-s \geq \delta}} \sup_{\substack{\mathbf{u} \in [0,1]^d \\ u_j \in [\varepsilon, 1-\varepsilon]}} \nu_{\mathbf{u}}^{\mathbf{x}_{[ns]+1: [nt]}} \left\{ \mathbf{w} \in [0,1]^d : w_j > 1 - \varepsilon/2 \right\}.
 \end{aligned}$$

We have that  $I'_{j,n,\delta,\varepsilon} = o_{\mathbb{P}}(1)$  as a consequence of Corollary 4.5.7. We shall now show that both  $I''_{j,n,\delta,\varepsilon}$  and  $I'''_{j,n,\delta,\varepsilon}$  converge almost surely to zero. To do so, it suffices to show that  $I''_{j,n,\delta,\varepsilon}$  and  $I'''_{j,n,\delta,\varepsilon}$  converge to zero conditionally on  $\mathbf{X}_1, \mathbf{X}_2, \dots$  for almost any sequence  $\mathbf{X}_1, \mathbf{X}_2, \dots$ . Concerning  $I''_{j,n,\delta,\varepsilon}$ , using Chebyshev's inequality and Condition 4.5.8, for almost any sequence  $\mathbf{X}_1, \mathbf{X}_2, \dots$ , conditionally on  $\mathbf{X}_1, \mathbf{X}_2, \dots$ , we obtain that

$$\begin{aligned}
 I''_{j,n,\delta,\varepsilon} &= \sup_{\substack{(s,t) \in \Lambda \\ t-s \geq \delta}} \sup_{u_j \in [\varepsilon, 1-\varepsilon]} \mathbb{P} \left\{ W_{j,u_j}^{\mathbf{x}_{[ns]+1: [nt]}} < \varepsilon/2 \mid \mathbf{x}_{[ns]+1: [nt]} \right\} \\
 &= \sup_{\substack{(s,t) \in \Lambda \\ t-s \geq \delta}} \sup_{u_j \in [\varepsilon, 1-\varepsilon]} \mathbb{P} \left\{ W_{j,u_j}^{\mathbf{x}_{[ns]+1: [nt]}} - u_j < \varepsilon/2 - u_j \mid \mathbf{x}_{[ns]+1: [nt]} \right\} \\
 &\leq \sup_{\substack{(s,t) \in \Lambda \\ t-s \geq \delta}} \sup_{u_j \in [\varepsilon, 1-\varepsilon]} \mathbb{P} \left\{ - \left| W_{j,u_j}^{\mathbf{x}_{[ns]+1: [nt]}} - u_j \right| \leq -u_j + \varepsilon/2 \mid \mathbf{x}_{[ns]+1: [nt]} \right\} \\
 &\leq \sup_{\substack{(s,t) \in \Lambda \\ t-s \geq \delta}} \sup_{u_j \in [\varepsilon, 1-\varepsilon]} \frac{\text{Var} \left( W_{j,u_j}^{\mathbf{x}_{[ns]+1: [nt]}} \mid \mathbf{x}_{[ns]+1: [nt]} \right)}{(u_j - \varepsilon/2)^2} \\
 &\leq \sup_{\substack{(s,t) \in \Lambda \\ t-s \geq \delta}} \sup_{u_j \in [\varepsilon, 1-\varepsilon]} \frac{a_{[nt]-[ns]}}{(u_j - \varepsilon/2)^2} \leq \sup_{\substack{(s,t) \in \Lambda \\ t-s \geq \delta}} a_{[nt]-[ns]} \sup_{u_j \in [\varepsilon, 1-\varepsilon]} \frac{1}{(u_j - \varepsilon/2)^2} \\
 &\leq \frac{4}{\varepsilon^2} \sup_{k \geq [n\delta]-1} a_k \rightarrow 0.
 \end{aligned}$$

Similarly, concerning  $I'''_{j,n,\delta,\varepsilon}$ , for almost any sequence  $\mathbf{X}_1, \mathbf{X}_2, \dots$ , conditionally on  $\mathbf{X}_1, \mathbf{X}_2, \dots$ , we obtain that

$$\begin{aligned}
 I'''_{j,n,\delta,\varepsilon} &= \sup_{\substack{(s,t) \in \Lambda \\ t-s \geq \delta}} \sup_{u_j \in [\varepsilon, 1-\varepsilon]} \mathbb{P} \left\{ W_{j,u_j}^{\mathbf{x}_{[ns]+1:[nt]}} > 1 - \varepsilon/2 \mid \mathbf{x}_{[ns]+1:[nt]} \right\} \\
 &= \sup_{\substack{(s,t) \in \Lambda \\ t-s \geq \delta}} \sup_{u_j \in [\varepsilon, 1-\varepsilon]} \mathbb{P} \left\{ W_{j,u_j}^{\mathbf{x}_{[ns]+1:[nt]}} - u_j > 1 - \varepsilon/2 - u_j \mid \mathbf{x}_{[ns]+1:[nt]} \right\} \\
 &\leq \sup_{\substack{(s,t) \in \Lambda \\ t-s \geq \delta}} \sup_{u_j \in [\varepsilon, 1-\varepsilon]} \mathbb{P} \left\{ \left| W_{j,u_j}^{\mathbf{x}_{[ns]+1:[nt]}} - u_j \right| \geq 1 - \varepsilon/2 - u_j \mid \mathbf{x}_{[ns]+1:[nt]} \right\} \\
 &\leq \sup_{\substack{(s,t) \in \Lambda \\ t-s \geq \delta}} \sup_{u_j \in [\varepsilon, 1-\varepsilon]} \frac{\text{Var} \left( W_{j,u_j}^{\mathbf{x}_{[ns]+1:[nt]}} \mid \mathbf{x}_{[ns]+1:[nt]} \right)}{(1 - \varepsilon/2 - u_j)^2} \\
 &\leq \sup_{\substack{(s,t) \in \Lambda \\ t-s \geq \delta}} \sup_{u_j \in [\varepsilon, 1-\varepsilon]} \frac{a_{[nt]-[ns]}}{(1 - \varepsilon/2 - u_j)^2} \\
 &\leq \sup_{\substack{(s,t) \in \Lambda \\ t-s \geq \delta}} a_{[nt]-[ns]} \sup_{u_j \in [\varepsilon, 1-\varepsilon]} \frac{1}{(1 - \varepsilon/2 - u_j)^2} \leq \frac{4}{\varepsilon^2} \sup_{k \geq [n\delta]-1} a_k \rightarrow 0.
 \end{aligned}$$

*Term  $J_{j,n,\delta,\varepsilon}$ :* Let  $\eta > 0$  and let us show that  $J_{j,n,\delta,\varepsilon} \leq \eta$  for  $n$  sufficiently large. For any  $\rho \in (0, 1)$ , from the triangle inequality and the fact that  $0 \leq \dot{C}_j \leq 1$ , we have that  $J_{j,n,\delta,\varepsilon}$  is smaller than

$$\begin{aligned}
 &\sup_{\substack{(s,t) \in \Lambda \\ t-s \geq \delta}} \sup_{u_j \in [\varepsilon, 1-\varepsilon]} \left| \int_{\{\mathbf{w} \in [0,1]^d : |\mathbf{w}-\mathbf{u}|_\infty \leq \rho\}} \{\dot{C}_j(\mathbf{w}) - \dot{C}_j(\mathbf{u})\} d\nu_{\mathbf{u}}^{\mathbf{x}_{[ns]+1:[nt]}}(\mathbf{w}) \right| \\
 &+ \sup_{\substack{(s,t) \in \Lambda \\ t-s \geq \delta}} \sup_{u_j \in [\varepsilon, 1-\varepsilon]} \left| \int_{\{\mathbf{w} \in [0,1]^d : |\mathbf{w}-\mathbf{u}|_\infty > \rho\}} \{\dot{C}_j(\mathbf{w}) - \dot{C}_j(\mathbf{u})\} d\nu_{\mathbf{u}}^{\mathbf{x}_{[ns]+1:[nt]}}(\mathbf{w}) \right| \\
 &\leq J'_{j,\varepsilon,\rho} + J''_{j,n,\delta,\rho},
 \end{aligned}$$

where

$$\begin{aligned}
 J'_{j,\varepsilon,\rho} &= \sup_{\substack{\mathbf{u} \in [0,1]^d \\ u_j \in [\varepsilon, 1-\varepsilon]}} \sup_{\substack{\mathbf{w} \in [0,1]^d \\ |\mathbf{w}-\mathbf{u}|_\infty \leq \rho}} \left| \dot{C}_j(\mathbf{w}) - \dot{C}_j(\mathbf{u}) \right|, \\
 J''_{j,n,\delta,\rho} &= \sup_{\substack{(s,t) \in \Lambda \\ t-s \geq \delta}} \sup_{\mathbf{u} \in [0,1]^d} \int_{[0,1]^d} \mathbf{1}\{|\mathbf{w}-\mathbf{u}|_\infty > \rho\} d\nu_{\mathbf{u}}^{\mathbf{x}_{[ns]+1:[nt]}}(\mathbf{w}).
 \end{aligned}$$

From Condition 4.2.8,  $\dot{C}_j$  is uniformly continuous on the set  $\{\mathbf{u} \in [0,1]^d : u_j \in [\varepsilon/2, 1 - \varepsilon/2]\}$ . We then choose  $\rho = \rho(\varepsilon, \eta) > 0$  sufficiently small such that

$$J'_{j,\varepsilon,\rho} = \sup_{\substack{\mathbf{u} \in [0,1]^d \\ u_j \in [\varepsilon, 1-\varepsilon]}} \sup_{\substack{\mathbf{w} \in [0,1]^d \\ |\mathbf{w}-\mathbf{u}|_\infty \leq \rho}} \left| \dot{C}_j(\mathbf{w}) - \dot{C}_j(\mathbf{u}) \right| \leq \frac{\eta}{2}. \tag{4.78}$$

As far as  $J''_{j,n,\delta,\rho}$  is concerned, using Chebyshev's inequality and Condition 4.5.8, for almost any sequence  $\mathbf{X}_1, \mathbf{X}_2, \dots$ , conditionally on  $\mathbf{X}_1, \mathbf{X}_2, \dots$ , we obtain that

$$\begin{aligned}
 J''_{j,n,\delta,\rho} &= \sup_{\substack{(s,t) \in \Lambda \\ t-s \geq \delta}} \sup_{\mathbf{u} \in [0,1]^d} \nu_{\mathbf{u}}^{\mathbf{x}_{[ns]+1:[nt]}} (\{\mathbf{w} \in [0,1]^d : |\mathbf{u} - \mathbf{w}|_\infty > \rho\}) \\
 &= \sup_{\substack{(s,t) \in \Lambda \\ t-s \geq \delta}} \sup_{\mathbf{u} \in [0,1]^d} \mathbb{P} \left[ \bigcup_{j=1}^d \left\{ |W_{j,u_j}^{\mathbf{x}_{[ns]+1:[nt]}} - u_j| > \rho \right\} \mid \mathbf{x}_{[ns]+1:[nt]} \right] \\
 &\leq \sum_{j=1}^d \sup_{\substack{(s,t) \in \Lambda \\ t-s \geq \delta}} \sup_{\mathbf{u} \in [0,1]^d} \mathbb{P} \left\{ |W_{j,u_j}^{\mathbf{x}_{[ns]+1:[nt]}} - u_j| > \rho \mid \mathbf{x}_{[ns]+1:[nt]} \right\} \\
 &\leq \sum_{j=1}^d \sup_{\substack{(s,t) \in \Lambda \\ t-s \geq \delta}} \sup_{\mathbf{u} \in [0,1]^d} \frac{\text{Var} \left( W_{j,u_j}^{\mathbf{x}_{[ns]+1:[nt]}} \mid \mathbf{x}_{[ns]+1:[nt]} \right)}{\rho^2} \\
 &\leq \frac{d}{\rho^2} \sup_{\substack{(s,t) \in \Lambda \\ t-s \geq \delta}} a_{[nt]-[ns]} \leq \frac{d}{\rho^2} \sup_{k \geq [n\delta]-1} a_k \rightarrow 0,
 \end{aligned}$$

which implies that, for  $n$  sufficiently large, with probability 1,  $J''_{j,n,\delta,\rho} \leq \eta/2$ . Using additionally (4.78), we obtain that  $J_{j,n,\delta,\varepsilon}$  converges almost surely to zero, which concludes the proof of (4.53). The proof of the analogous result for  $\dot{C}_{j,k:l}^{\nabla,\nu}$  in (4.48) is almost identical.  $\blacksquare$

**Proof of Proposition 4.5.10.** Fix  $j \in \{1, \dots, d\}$ ,  $\delta \in (0, 1)$  and  $\varepsilon \in (0, 1/2)$ . From (4.50), we have that, for any  $(s, t) \in \Lambda$  and  $\mathbf{u} \in [0, 1]^d$ ,

$$\dot{C}_{j,[ns]+1:[nt],m_{[nt]-[ns]}}^{\text{Bern}}(\mathbf{u}) = \int_{[0,1]^d} \dot{C}_{j,[ns]+1:[nt],1/m_{[nt]-[ns]},0}^{\nabla}(\mathbf{w}) d\tilde{\mu}_{j,m_{[nt]-[ns]},\mathbf{u}}(\mathbf{w}),$$

where  $\dot{C}_{j,k:l,1/m,0}^{\nabla}$  is defined in (4.38) and, for any  $m \geq 2$ ,  $\tilde{\mu}_{j,m,\mathbf{u}}$  is the law of the random vector  $(\tilde{S}_{m,1,u_1}/m, \dots, \tilde{S}_{m,d,u_d}/m)$  whose components are independent such that, for  $i \in \{1, \dots, d\} \setminus \{j\}$ ,  $\tilde{S}_{m,i,u_i}$  is Binomial( $m, u_i$ ) while  $\tilde{S}_{m,j,u_j}$  is Binomial( $m-1, u_j$ ). It follows that, for any  $(s, t) \in \Lambda$  and  $\mathbf{u} \in [0, 1]^d$ ,

$$\dot{C}_{j,[ns]+1:[nt],m_{[nt]-[ns]}}^{\text{Bern}}(\mathbf{u}) = \int_{\mathcal{W}_{j,n,s,t}} \dot{C}_{j,[ns]+1:[nt],1/m_{[nt]-[ns]},0}^{\nabla}(\mathbf{w}) d\tilde{\mu}_{j,m_{[nt]-[ns]},\mathbf{u}}(\mathbf{w}), \quad (4.79)$$

where  $\mathcal{W}_{j,n,s,t} = \{\mathbf{w} \in [0, 1]^d : w_j \leq 1 - 1/m_{[nt]-[ns]}\}$ . For the sake of a more compact notation, from now on, we shall write  $m_{s,t}$  for  $m_{[nt]-[ns]}$ ,  $(s, t) \in \Lambda$ . From the triangle inequality, the

left-hand side of (4.54) is smaller than  $I_{j,n,\delta,\varepsilon} + J_{j,n,\delta,\varepsilon}$ , where

$$I_{j,n,\delta,\varepsilon} = \sup_{\substack{(s,t) \in \Lambda \\ t-s \geq \delta}} \sup_{\substack{\mathbf{u} \in [0,1]^d \\ u_j \in [\varepsilon, 1-\varepsilon]}} \left| \int_{\mathcal{W}_{j,n,s,t}} \left\{ \dot{C}_{j,[ns]+1:[nt],1/m_{s,t},0}^\nabla(\mathbf{w}) - \dot{C}_j(\mathbf{w}) \right\} d\tilde{\mu}_{j,m_{s,t},\mathbf{u}}(\mathbf{w}) \right|,$$

$$J_{j,n,\delta,\varepsilon} = \sup_{\substack{(s,t) \in \Lambda \\ t-s \geq \delta}} \sup_{\substack{\mathbf{u} \in [0,1]^d \\ u_j \in [\varepsilon, 1-\varepsilon]}} \left| \int_{[0,1]^d} \dot{C}_j(\mathbf{w}) d\tilde{\mu}_{j,m_{s,t},\mathbf{u}}(\mathbf{w}) - \dot{C}_j(\mathbf{u}) \right|.$$

For any  $n \in \mathbb{N}$ ,  $\mathbf{x} \in (\mathbb{R}^d)^n$  and  $\mathbf{u} \in [0,1]^d$ , let  $\nu_{\mathbf{u}}^{\mathbf{x}} = \tilde{\mu}_{j, \lfloor Ln^\theta \rfloor \vee 2, \mathbf{u}}$ . With this notation, Condition 4.5.8 holds for the considered smoothing distributions and it can be verified that  $J_{j,n,\delta,\varepsilon} = o_{\mathbb{P}}(1)$  by proceeding exactly as in the proof of Proposition 4.5.9 for the analogous term. It thus remain to show that  $I_{j,n,\delta,\varepsilon} = o_{\mathbb{P}}(1)$ .

From the triangle inequality, we have that  $I_{j,n,\delta,\varepsilon}$  is smaller than

$$\begin{aligned} & \sup_{\substack{(s,t) \in \Lambda \\ t-s \geq \delta}} \sup_{\substack{\mathbf{u} \in [0,1]^d \\ u_j \in [\varepsilon, 1-\varepsilon]}} \left| \int_{\{\mathbf{w} \in \mathcal{W}_{j,n,s,t} : w_j \in [\varepsilon/2, 1-\varepsilon/2]\}} \left\{ \dot{C}_{j,[ns]+1:[nt],1/m_{s,t},0}^\nabla(\mathbf{w}) - \dot{C}_j(\mathbf{w}) \right\} d\tilde{\mu}_{j,m_{s,t},\mathbf{u}}(\mathbf{w}) \right| \\ & + \sup_{\substack{(s,t) \in \Lambda \\ t-s \geq \delta}} \sup_{\substack{\mathbf{u} \in [0,1]^d \\ u_j \in [\varepsilon, 1-\varepsilon]}} \left| \int_{\{\mathbf{w} \in \mathcal{W}_{j,n,s,t} : w_j < \varepsilon/2\}} \left\{ \dot{C}_{j,[ns]+1:[nt],1/m_{s,t},0}^\nabla(\mathbf{w}) - \dot{C}_j(\mathbf{w}) \right\} d\tilde{\mu}_{j,m_{s,t},\mathbf{u}}(\mathbf{w}) \right| \\ & + \sup_{\substack{(s,t) \in \Lambda \\ t-s \geq \delta}} \sup_{\substack{\mathbf{u} \in [0,1]^d \\ u_j \in [\varepsilon, 1-\varepsilon]}} \left| \int_{\{\mathbf{w} \in \mathcal{W}_{j,n,s,t} : w_j > 1-\varepsilon/2\}} \left\{ \dot{C}_{j,[ns]+1:[nt],1/m_{s,t},0}^\nabla(\mathbf{w}) - \dot{C}_j(\mathbf{w}) \right\} d\tilde{\mu}_{j,m_{s,t},\mathbf{u}}(\mathbf{w}) \right| \\ & \leq I'_{j,n,\delta,\varepsilon} + M_n I''_{j,n,\delta,\varepsilon} + M_n I'''_{j,n,\delta,\varepsilon}, \end{aligned}$$

where

$$I'_{j,n,\delta,\varepsilon} = \sup_{\substack{(s,t) \in \Lambda \\ t-s \geq \delta}} \sup_{\substack{\mathbf{w} \in [0,1]^d \\ w_j \in [\varepsilon/2, 1-\varepsilon/2]}} \left| \dot{C}_{j,[ns]+1:[nt],1/m_{s,t},0}^\nabla(\mathbf{w}) - \dot{C}_j(\mathbf{w}) \right|,$$

$$I''_{j,n,\delta,\varepsilon} = \sup_{\substack{(s,t) \in \Lambda \\ t-s \geq \delta}} \sup_{\substack{\mathbf{u} \in [0,1]^d \\ u_j \in [\varepsilon, 1-\varepsilon]}} \tilde{\mu}_{j,m_{s,t},\mathbf{u}} \left\{ \mathbf{w} \in [0,1]^d : w_j < \varepsilon/2 \right\},$$

$$I'''_{j,n,\delta,\varepsilon} = \sup_{\substack{(s,t) \in \Lambda \\ t-s \geq \delta}} \sup_{\substack{\mathbf{u} \in [0,1]^d \\ u_j \in [\varepsilon, 1-\varepsilon]}} \tilde{\mu}_{j,m_{s,t},\mathbf{u}} \left\{ \mathbf{w} \in [0,1]^d : w_j > 1 - \varepsilon/2 \right\},$$

$$M_n = \sup_{(s,t) \in \Lambda} \sup_{\mathbf{w} \in \mathcal{W}_{j,n,s,t}} \left| \dot{C}_{j,[ns]+1:[nt],1/m_{s,t},0}^\nabla(\mathbf{w}) \right|.$$

Since the conditions of the proposition imply that Condition 4.5.4 holds with  $h(\mathbf{x}) = 1/(\lfloor Ln^\theta \rfloor \vee 2)$  and  $h'(\mathbf{x}) = 0$  for all  $n \in \mathbb{N}$  and  $\mathbf{x} \in (\mathbb{R}^d)^n$ , we have that  $I'_{j,n,\delta,\varepsilon} = o_{\mathbb{P}}(1)$  as a consequence of Corollary 4.5.7. Also, given that Condition 4.5.8 holds for the considered smoothing distributions, it can be verified that  $I''_{j,n,\delta,\varepsilon}$  and  $I'''_{j,n,\delta,\varepsilon}$  converge almost surely to zero by proceeding exactly as in the proof of Proposition 4.5.9 for the analogous terms. To complete the proof

of (4.54), it suffices to show that, there exists a constant  $\zeta > 0$  such that, for any  $n \in \mathbb{N}$ ,  $M_n < \zeta$  almost surely.

Fix  $n \in \mathbb{N}$ . From the adopted conventions, we have that  $\dot{C}_{j, [ns]+1: [nt], 1/m_{s,t}, 0}^\nabla = 0$  for all  $(s, t) \in \Lambda$  such that  $[ns] = [nt]$ . Fix  $(s, t) \in \Lambda$  such that  $[ns] < [nt]$  and let  $p = [nt] - [ns]$ . The empirical copula  $C_{[ns]+1: [nt]}$ , generically defined in (4.2), is a multivariate d.f. whose  $d$  univariate margins, under Condition 4.2.3, are all equal to  $G_{[ns]+1: [nt]}$ , where  $G_{[ns]+1: [nt]}(u) = [pu]/p$ ,  $u \in [0, 1]$ . As a consequence of a well-known property of multivariate d.f.s (see, e.g., Durante and Sempi, 2015, Lemma 1.2.14), we have that

$$|C_{[ns]+1: [nt]}(\mathbf{u}) - C_{[ns]+1: [nt]}(\mathbf{v})| \leq \sum_{j=1}^d |G_{[ns]+1: [nt]}(u_j) - G_{[ns]+1: [nt]}(v_j)|$$

for all  $\mathbf{u}, \mathbf{v} \in [0, 1]^d$ . We then obtain that, for any  $\mathbf{u} \in \mathcal{W}_{j, n, s, t}$ ,

$$|C_{[ns]+1: [nt]}(\mathbf{u} + \mathbf{e}_j/m_{s,t}) - C_{[ns]+1: [nt]}(\mathbf{u})| \leq |G_{[ns]+1: [nt]}(u_j + 1/m_{s,t}) - G_{[ns]+1: [nt]}(u_j)|,$$

which implies that

$$\begin{aligned} \left| \dot{C}_{j, [ns]+1: [nt]}(\mathbf{u}) \right| &\leq \frac{|G_{[ns]+1: [nt]}(u_j + 1/m_{s,t}) - G_{[ns]+1: [nt]}(u_j)|}{1/m_{s,t}} \\ &= m_{s,t} \left\{ \frac{[p(u_j + 1/m_{s,t})]}{p} - \frac{[pu_j]}{p} \right\} \\ &\leq m_{s,t} \left\{ \frac{p(u_j + 1/m_{s,t})}{p} - \frac{pu_j - 1}{p} \right\} \\ &\leq m_{s,t} \left( \frac{1}{m_{s,t}} + \frac{1}{p} \right) \leq 1 + \frac{m_{s,t}}{p} = 1 + \frac{[Lp^\theta] \vee 2}{p} \\ &\leq 1 + Lp^{\theta-1} \vee (2/p) \leq 1 + L \vee 2, \end{aligned}$$

which completes the proof of (4.54). The fact that (4.55) holds is finally an immediate consequence of the previous centered display and (4.79).  $\blacksquare$

# Chapter 5

## Conclusion and some future research directions

5.1	Conclusion	133
5.2	Some future research directions	134
5.2.1	Application to sequential change-point detection	134
5.2.2	Application to high-dimensional factor copula models	135

### 5.1 Conclusion

In this work, we proposed and studied a broad class of smooth, possibly data-adaptive empirical copulas that contain the empirical beta copula. Within this class, the estimator  $C_{1:n}^{\bar{B}, C_{1:n}^{\text{beta}}}$  generically defined in (3.18) is uniformly better than the empirical beta copula for all the considered data-generating models. Specifically,  $C_{1:n}^{\bar{B}, C_{1:n}^{\text{beta}}}$  uses the empirical beta copula as survival copula of the smoothing distribution and scaled beta-binomial smoothing survival margins with dispersion parameter  $\rho = 4$ . From Corollary 3.3.8, the latter is a genuine copula in the absence of ties in the component samples. From a theoretical perspective, for the purpose of applying the studied class of smooth estimators to change-point detection, we provided conditions under which the corresponding smooth sequential empirical copula processes converge weakly.

As explained in previous chapters, in order to use such estimators in inference procedures, one typically needs to rely on resampling techniques. In this work, we investigated two resampling schemes for the class of studied smooth estimators. The first scheme is an extension of the smooth bootstrap in Kiriliouk et al. (2021) that involves drawing samples from the estimators. Asymptotic validity of this smooth bootstrap was established and numerical results show that it provides better-behaved inference procedures than the one considered in Kiriliouk et al. (2021) as it copes better with stronger dependence. Unfortunately, such smooth bootstraps cannot be

used anymore in the time series setting. The second resampling scheme that we investigated is a smooth extension of the sequential dependent multiplier bootstrap of [Bücher and Kojadinovic \(2016\)](#), which can be used in the sequential and time series setting. Asymptotic validity of this resampling scheme was established and numerical results show that when it is used in change-point detection, the smooth change-point detection test is likely to be more sensitive to early or late changes than their non-smooth counterparts since, as already mentioned, smooth estimators are likely to be more accurate than the empirical copula when computed from small subsets of observations.

In connection with the multiplier bootstrap, two classes of data-adaptive smooth estimators of the first-order partial derivatives of the copula were also proposed. Weak consistency of these estimators was studied and numerical results show that a smooth data-adaptive estimator of the first-order partial derivatives of the copula outperforms substantially, among others, the Bernstein estimator studied in [Janssen et al. \(2016\)](#). These results may be of independent interest since such estimators have applications in mean and quantile regression as they lead to estimators of the conditional distribution function. Furthermore, they can be applied to *factor-copula models* as will be explained in the next section.

## 5.2 Some future research directions

In this section, we discuss two directions of extension that could lead to future projects.

### 5.2.1 Application to sequential change-point detection

In Section 4.4.6 of Chapter 4, we explain how the smooth sequential empirical copula process  $\mathbb{C}_n^\nu$  in Chapter 3 can be applied to *offline change-point detection*, where the whole time series is accessible at once; see also Example 2.3.19. This setup is different from what is called *sequential change-point detection* (also referred to as *online change-point detection*), where new observations arrive sequentially and we need to make a decision at each step whether a change has already occurred (see, e.g., [Kojadinovic and Verdier, 2021](#); [Holmes and Kojadinovic, 2021](#)).

The extension of  $\mathbb{C}_n^\nu$  to sequential change-point detection can be built up on [Kojadinovic and Verdier \(2021\)](#). Specifically, let  $\mathbf{X}_1, \dots, \mathbf{X}_m, m > 1$ , be a stretch from a  $d$ -dimensional stationary time series of continuous random vectors with unknown contemporary copula  $C$  and marginal d.f.s  $F_1, \dots, F_d$ . These available observations will be referred to as the *learning sample*. New observations  $\mathbf{X}_{m+1}, \mathbf{X}_{m+2}, \dots$  arrive sequentially and we wish to issue an alarm as soon as possible if the contemporary copula is not equal to  $C$  anymore. If there is no evidence of a change in copula, the monitoring stops after the arrival of observation  $\mathbf{X}_n$  for some  $n > m$ .

The null hypothesis can be formally stated as follows:

$H_0$ :  $\mathbf{X}_1, \dots, \mathbf{X}_m, \mathbf{X}_{m+1}, \dots, \mathbf{X}_n$  is a stretch from a stationary time series with contemporary copula  $C$  and marginal d.f.s  $F_1, \dots, F_d$ .

An alternative hypothesis of particular interest is:

$H_1$ :  $\exists$  distinct  $C_1$  and  $C_2$ , and  $k^* \in \{m, \dots, n-1\}$  such that  $\mathbf{X}_1, \dots, \mathbf{X}_{k^*}$  is a stretch from a stationary time series with contemporary copula  $C_1$  and  $\mathbf{X}_{k^*+1}, \dots, \mathbf{X}_n$  is a stretch from a stationary time series with contemporary copula  $C_2 \neq C_1$ .

Test statistics can be adapted from [Kojadinovic and Verdier \(2021\)](#). For instance, for any  $k \in \{m+1, \dots, n\}$ , let

$$R_{m,q}^\nu(k) = \max_{m \leq j \leq k-1} \frac{j(k-j)}{m^{3/2}q(j/m, k/m)} \sup_{\mathbf{u} \in [0,1]^d} |C_{1:j}^\nu(\mathbf{u}) - C_{j+1:k}^\nu(\mathbf{u})|, \quad (5.1)$$

where  $C_{k:l}^\nu$  is the smooth estimator defined in (4.2) and  $q$  is some weight function. A reason for using a smooth estimator rather than the classical empirical copula is that the former can be a substantially better estimator than the latter in small samples. In this case, the tests based on the smooth estimator should be more sensitive to early changes than the ones based on the empirical copula. Under  $H_0$  in (5.2.1),  $R_{m,q}^\nu(k)$  can be written as a functional of the smooth empirical process  $\mathbb{C}_n^\nu$  in (4.10) and the multiplier bootstraps in Section 4.4 of Chapter 4 can be applied to obtain replicates of the test statistic and approximate  $p$ -values.

## 5.2.2 Application to high-dimensional factor copula models

Let us first briefly review the one-factor copula models proposed in [Kurpskii and Joe \(2020\)](#). Let  $V$  be a standard uniform random variable and let  $(U_j)_{j \in \mathbb{N}}$  be a sequence of standard uniform random variables that are conditionally independent given  $V$ . In this construction,  $V$  is a latent random variable. For any  $d \in \mathbb{N}$ , denote by  $C_{U_1, \dots, U_d}$  the copula of  $(U_1, \dots, U_d)$  and by  $C_{U_1, \dots, U_d, V}$  the copula of  $(U_1, \dots, U_d, V)$ . Also, for any  $j \in \mathbb{N}$ , define the function

$$\dot{C}_j(u, v) = \frac{\partial C_{U_j, V}(u, v)}{\partial v}, \quad (u, v) \in [0, 1] \times (0, 1). \quad (5.2)$$

For any  $d \in \mathbb{N}$  and  $\mathbf{u} = (u_1, \dots, u_d) \in [0, 1]^d$ , we can then write

$$\begin{aligned} C_{U_1, \dots, U_d}(\mathbf{u}) &= \mathbb{P}(U_1 \leq u_1, \dots, U_d \leq u_d) = \int_0^1 \mathbb{P}(U_1 \leq u_1, \dots, U_d \leq u_d \mid V = v) dv \\ &= \int_0^1 \prod_{j=1}^d \mathbb{P}(U_j \leq u_j \mid V = v) dv = \int_0^1 \prod_{j=1}^d \dot{C}_j(u_j, v) dv. \end{aligned} \quad (5.3)$$

Hence, as one can see, for any  $d \in \mathbb{N}$ , the distribution of  $(U_1, \dots, U_d)$  depends only on the copulas  $C_{U_j, V}$  linking  $U_j$  and  $V$ , for  $j \in \{1, \dots, d\}$ . However, since the latent variable  $V$  is



unobservable, Kurpskii and Joe (2020) proposed to use the proxy variable  $\bar{U}_d = \frac{1}{d} \sum_{j=1}^d U_j$  when the dimension is large. According to Proposition 5 of Kurpskii and Joe (2020), under regularity conditions, for any  $k \in \mathbb{N}$  and  $\mathbf{u} \in [0, 1]^{k+1}$ ,

$$C_{U_1, \dots, U_k, \bar{U}_d}(\mathbf{u}) - C_{U_1, \dots, U_k, V}(\mathbf{u}) \rightarrow 0, \quad \text{as } d \rightarrow \infty,$$

where  $C_{U_1, \dots, U_k, \bar{U}_d}$  denotes the copula of  $(U_1, \dots, U_k, \bar{U}_d)$ . As a consequence, for any  $j \in \{1, \dots, k\}$  and  $\mathbf{u} \in [0, 1]^2$ ,  $C_{U_j, \bar{U}_d}(\mathbf{u}) - C_{U_j, V}(\mathbf{u}) \rightarrow 0$ , as  $d \rightarrow \infty$ , where  $C_{U_j, \bar{U}_d}$  (resp.  $C_{U_j, V}$ ) denotes the copula of  $(U_j, \bar{U}_d)$  (resp.  $(U_j, V)$ ).

The conjecture is that, when  $d$  is a large,  $C_{U_1, \dots, U_d}(\mathbf{u}) \approx \tilde{C}_{U_1, \dots, U_d}(\mathbf{u})$  for all  $\mathbf{u} \in [0, 1]^d$ , where

$$\tilde{C}_{U_1, \dots, U_d}(\mathbf{u}) = \int_0^1 \prod_{j=1}^d \dot{C}_{j, \bar{U}_d}(u_j, v) dv, \quad \mathbf{u} \in [0, 1]^d, \quad (5.4)$$

$$\dot{C}_{j, \bar{U}_d}(u, v) = \frac{\partial C_{U_j, \bar{U}_d}(u, v)}{\partial v}, \quad (u, v) \in [0, 1] \times (0, 1), \quad j \in \{1, \dots, d\}. \quad (5.5)$$

Then, we can use the ‘‘plug-in’’ principle to propose an estimator of  $C_{U_1, \dots, U_d}$  in (5.3). Specifically, for any  $j \in \{1, \dots, d\}$ , let  $\hat{C}_{j, n, \bar{U}_d}$  be an estimator of  $\dot{C}_{j, \bar{U}_d}$  in (5.5). A natural estimator of  $C_{U_1, \dots, U_d}$  in (5.3) is then

$$\hat{C}_{U_1, \dots, U_d}(\mathbf{u}) = \int_0^1 \prod_{j=1}^d \hat{C}_{j, n, \bar{U}_d}(u_j, v) dv, \quad \mathbf{u} \in [0, 1]^d. \quad (5.6)$$

For the choice of  $\hat{C}_{j, n, \bar{U}_d}$ , we can use the two classes of smooth estimators of the first-order partial derivatives of the copula studied in Section 4.5 of Chapter 4. Furthermore, the consistency of  $\hat{C}_{U_1, \dots, U_d}$  can be studied and numerical experiments could be carried out to measure its finite-sample performance.

Moreover, note that the aforementioned ideas could even be extended to one-factor models with residual dependence (see Section 2.3 of Kurpskii and Joe (2020)), to *oblique factor copulas* (see Section 3.2 of Kurpskii and Joe (2020)) or to *nested copula models* proposed and studied in Kurpskii and Joe (2015).

# Bibliography

- Ali, M. M., N. Mikhail, and M. Haq (1978). A class of bivariate distributions including the bivariate logistic. *Journal of Multivariate Analysis* 8(3), 405–412.
- Berghaus, B., A. Bücher, and S. Volgushev (2017). Weak convergence of the empirical copula process with respect to weighted metrics. *Bernoulli* 23(1), 743–772.
- Bouezmarni, T., A. El Gouch, and A. Taamouti (2013). Bernstein estimator for unbounded copula densities. *Statistics & Risk Modeling* 30(4), 343–360.
- Bücher, A. (2015). A note on weak convergence of the sequential multivariate empirical process under strong mixing. *Journal of Theoretical Probability* 28(3), 1028–1037.
- Bücher, A. and H. Dette (2010). A note on bootstrap approximations for the empirical copula process. *Statistics and Probability Letters* 80(23–24), 1925–1932.
- Bücher, A. and I. Kojadinovic (2016). A dependent multiplier bootstrap for the sequential empirical copula process under strong mixing. *Bernoulli* 22(2), 927–968.
- Bücher, A. and I. Kojadinovic (2019). A note on conditional versus joint unconditional weak convergence in bootstrap consistency results. *Journal of Theoretical Probability* 32, 1145–1165.
- Bücher, A., I. Kojadinovic, T. Rohmer, and J. Segers (2014). Detecting changes in cross-sectional dependence in multivariate time series. *Journal of Multivariate Analysis* 132, 111–128.
- Bücher, A. and M. Ruppert (2013). Consistent testing for a constant copula under strong mixing based on the tapered block multiplier technique. *Journal of Multivariate Analysis* 116, 208–229.
- Bücher, A. and J. Segers (2014). Extreme value copula estimation based on block maxima of a multivariate stationary time series. *Extremes* 17(3), 495–528.
- Bücher, A. and S. Volgushev (2013). Empirical and sequential empirical copula processes under serial dependence. *Journal of Multivariate Analysis* 119, 61–70.
- Bühlmann, P. (1993). *The blockwise bootstrap in time series and empirical processes*. Ph. D. thesis, ETH Zürich. Diss. ETH No. 10354.

- Chen, X. and Y. Fan (1999). Consistent hypothesis testing in semiparametric and nonparametric models for econometric time series. *Journal of Econometrics* 91, 373–401.
- Chen, X. and Y. Fan (2006). Estimation and model selection of semiparametric copula-based multivariate dynamic models under copula misspecification. *Journal of Econometrics* 135, 125–154.
- Clayton, D. (1978). A model for association in bivariate life tables and its application in epidemiological studies of familial tendency in chronic disease incidence. *Biometrika* 65(1), 141–151.
- Deheuvels, P. (1979). La fonction de dépendance empirique et ses propriétés: un test non paramétrique d'indépendance. *Acad. Roy. Belg. Bull. Cl. Sci. 5th Ser.* 65, 274–292.
- Doukhan, P., J. Fermanian, and G. Lang (2009). An empirical central limit theorem with applications to copulas under weak dependence. *Statistical Inference for Stochastic Processes* 12(1), 65–87.
- Durante, F. and C. Sempì (2010). Copula theory: An introduction. In P. Jaworski, F. Durante, W. Härdle, and W. Rychlik (Eds.), *Copula theory and its applications (Warsaw, 2009)*, Lecture Notes in Statistics, pp. 3–32. Springer-Verlag.
- Durante, F. and C. Sempì (2015). *Principles of Copula Theory*. Boca Raton, Florida: CRC Press.
- Feller, W. (1948). On the Kolmogorov-Smirnov Limit Theorems for Empirical Distributions. *The Annals of Mathematical Statistics* 19(2), 177 – 189.
- Fermanian, J. and O. Scaillet (2003). Nonparametric estimation of copulas for time series. *FAME Research Paper* (57).
- Fermanian, J.-D., D. Radulovic, and M. Wegkamp (2004). Weak convergence of empirical copula processes. *Bernoulli* 10(5), 847–860.
- Gänssler, P. and W. Stute (1987). *Seminar on empirical processes*. DMV Seminar 9. Basel: Birkhäuser.
- Genest, C. (1987). Frank's family of bivariate distributions. *Biometrika* 74(3), 549–555.
- Genest, C., K. Ghoudi, and L.-P. Rivest (1995). A semiparametric estimation procedure of dependence parameters in multivariate families of distributions. *Biometrika* 82, 543–552.
- Genest, C., J. G. Nešlehová, and B. Rémillard (2017). Asymptotic behavior of the empirical multilinear copula process under broad conditions. *Journal of Multivariate Analysis* 20, 82–110.
- Genest, C. and J. Nešlehová (2007). A primer on copulas for count data. *The Astin Bulletin* 37, 475–515.

- Genest, C., J. Nešlehová, and J.-F. Quessy (2012). Tests of symmetry for bivariate copulas. *Annals of the Institute of Statistical Mathematics* 64, 811–834.
- Genest, C. and B. Rémillard (2004). Tests of independence and randomness based on the empirical copula process. *Test* 13(2), 335–369.
- Gribkova, S. and O. Lopez (2015). Non-parametric copula estimation under bivariate censoring. *Scandinavian Journal of Statistics* 42(4), 925–946.
- Gumbel, E. (1961). Bivariate logistic distributions. *Journal of the American Statistical Association* 56(294), 335–349.
- Hofert, M., I. Kojadinovic, M. Mächler, and J. Yan (2022). *copula: Multivariate dependence with copulas*. R package version 1.1-0.
- Hofert, M., I. Kojadinovic, M. Mächler, and J. Yan (2018). *Elements of copula modeling with R*. Springer.
- Holmes, M. and I. Kojadinovic (2021). Open-end nonparametric sequential change-point detection based on the retrospective cusum statistic. *Electron. J. Statist.*.
- Holmes, M., I. Kojadinovic, and J.-F. Quessy (2013). Nonparametric tests for change-point detection à la Gombay and Horváth. *Journal of Multivariate Analysis* 115, 16–32.
- Hougaard, P. (1986). A class of multivariate failure time distributions. *Biometrika* 73(3), 671–678.
- Janssen, P., J. Swanepoel, and N. Veraverbeke (2012). Large sample behavior of the Bernstein copula estimator. *Journal of Statistical Planning and Inference* 142, 1189–1197.
- Janssen, P., J. Swanepoel, and N. Veraverbeke (2016). Bernstein estimation for a copula derivative with application to conditional distribution and regression functionals. *Test* 25, 351–374.
- Joe, H. (1997). *Multivariate models and dependence concepts*. London: Chapman and Hall.
- Joe, H. (2005). Asymptotic efficiency of the two-stage estimation method for copula-based models. *Journal of Multivariate Analysis* 94, 401–419.
- Joe, H. (2015). *Dependence Modeling with Copulas*. CRC Press.
- Joe, H. and J. J. Xu (1996). The estimation method of inference functions for margins for multivariate models. Technical report, Department of Statistics, University of British Columbia.
- Kendall, M. G. (1938, 06). A New Measure Of Rank Correlation. *Biometrika* 30(1-2), 81–93.
- Kiriliouk, A., J. Segers, and H. Tsukahara (2021). *Resampling Procedures with Empirical Beta Copulas*, pp. 27–53. Springer Singapore.
- Klassen, C. A. J. and J. A. Wellner (1997). Efficient estimation in the bivariate normal copula model: Normal marginals are least favourable. *Bernoulli* 3, 55–77.

- Kojadinovic, I., J. Segers, and J. Yan (2011). Large-sample tests of extreme-value dependence for multivariate copulas. *The Canadian Journal of Statistics* 39(4), 703–720.
- Kojadinovic, I. and K. Stemikovskaya (2019). Subsampling (weighted smooth) empirical copula processes. *Journal of Multivariate Analysis* 173, 704–723.
- Kojadinovic, I. and G. Verdier (2021). Nonparametric sequential change-point detection for multivariate time series based on empirical distribution functions. *Electron. J. Statist.* 15(1), 773–829.
- Kojadinovic, I. and J. Yan (2010). A nonparametric test of exchangeability for bivariate extreme-value copulas. Technical report 3, Department of Statistics, The University of Connecticut.
- Kojadinovic, I. and B. Yi (2022). A class of smooth, possibly data-adaptive nonparametric copula estimators containing the empirical beta copula. *Journal of Multivariate Analysis*, in press. <http://arxiv.org/abs/2106.10726>.
- Kojadinovic, I. and B. Yi (2023). Resampling techniques for a class of smooth, possibly data-adaptive empirical copulas. *preprint*. <https://arxiv.org/abs/2301.05495>.
- Kosorok, M. (2008). *Introduction to empirical processes and semiparametric inference*. New York: Springer.
- Krupskii, P. (2017). Copula-based measures of reflection and permutation asymmetry and statistical tests. *Statistical Papers* 58(4), 1165–1187.
- Krupskii, P. and H. Joe (2015). Structured factor copula models: Theory, inference and computation. *Journal of Multivariate Analysis* 138, 53–73.
- Künsch, H. (1989). The jackknife and the bootstrap for general stationary observations. *The Annals of Statistics* 17(3), 1217–1241.
- Kurpskii, P. and H. Joe (2013). Factor copula models for multivariate data. *Journal of Multivariate Analysis* 120, 85–101.
- Kurpskii, P. and H. Joe (2020). Approximate likelihood with proxy variables for parameter estimation in high-dimensional factor copula models.
- Nelsen, R. (2006). *An introduction to copulas*. New-York: Springer. Second edition.
- Philipp, W. and L. Pinzur (1980). Almost sure approximation theorems for the multivariate empirical process. *Probability Theory and Related Fields* 54, 1–13.
- R Core Team (2022). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.
- Rémillard, B. and O. Scaillet (2009). Testing for equality between two copulas. *Journal of Multivariate Analysis* 100(3), 377–386.

- Rosenblatt, M. (1952). Remarks on a Multivariate Transformation. *The Annals of Mathematical Statistics* 23(3), 470 – 472.
- Rüschendorf, L. (1976). Asymptotic distributions of multivariate rank order statistics. *The Annals of Statistics* 4, 912–923.
- Sancetta, A. and S. Satchell (2004). The Bernstein copula and its applications to modeling and approximations of multivariate distributions. *Econometric Theory* 20, 535–562.
- Scaillet, O. (2005). A Kolmogorov-Smirnov type test for positive quadrant dependence. *The Canadian Journal of Statistics* 33, 415–427.
- Segers, J. (2012). Asymptotics of empirical copula processes under nonrestrictive smoothness assumptions. *Bernoulli* 18, 764–782.
- Segers, J., M. Sibuya, and H. Tsukahara (2017). The empirical beta copula. *Journal of Multivariate Analysis* 155, 35–51.
- Shaked, M. and J. Shanthikumar (2007). *Stochastic orders*. Springer Series in Statistics. Springer.
- Shih, J. and T. Louis (1995). Inferences on the association parameter in copula models for bivariate survival data. *Biometrics* 51(4), 1384–1399.
- Shorack, G. and J. A. Wellner (1986). *Empirical processes with applications to statistics*. Wiley Series in Probability and Mathematical Statistics. New York: John Wiley and Sons.
- Sklar, A. (1959). Fonctions de répartition à  $n$  dimensions et leurs marges. *Publications de l'Institut de Statistique de l'Université de Paris* 8, 229–231.
- Spearman, C. (1904). The proof and measurement of association between two things. *The American Journal of Psychology* 15(1), 72–101.
- Tsukahara, H. (2005). Semiparametric estimation in copula models. *The Canadian Journal of Statistics* 33(3), 357–375.
- van der Vaart, A. (1998). *Asymptotic statistics*. Cambridge University Press.
- van der Vaart, A. and J. A. Wellner (2000). *Weak convergence and empirical processes*. New York: Springer. Second edition.
- van der Vaart, A. and J. A. Wellner (2007). Empirical processes indexed by estimated functions. In *Asymptotics: Particles, Processes and Inverse Problems*, pp. 234–252. Institute of Mathematical Statistics.
- Wolodzko, T. (2020). *extraDistr: Additional univariate and multivariate distributions*. R package version 1.9.1.