Modeling the dependence of pre-asymptotic extremes

Simon Chatelain

N° d'ordre NNT : 2019LYSE1267

THÈSE DE DOCTORAT DE L'UNIVERSITÉ DE LYON
opérée au sein de
l'Université Claude Bernard Lyon 1

École Doctorale ED512
InfoMaths

Discipline : Mathématiques

Soutenue publiquement le 17 décembre 2019, par :
Simon Chatelain

Modélisation de la dépendance entre pré-extrêmes

Devant le jury composé de :

John Einmahl	Professeur, Tilburg University	Rapporteur
Jean-David Fermanian	Professeur, ENSAE Paris	Rapporteur
Clément Dombry	Professeur, Université de Franche-Comté	Examinateur
Sebastian Engelke	Professeur assistant, Université de Genève	Examinateur
Véronique Maume-Deschamps	Professeure, Université de Lyon 1	Examinatrice
Russell Steele	Professeur associé, Université McGill	Examinateur
Anne-Laure Fougères	Professeure, Université de Lyon 1	Directrice de thèse
Johanna G. Nešlehová	Professeure associée, Université McGill	Directrice de thèse
Stanislav Volgushev	Professeur assistant, University of Toronto	Invité

Après avis des rapporteurs :

John Einmahl	Professeur, Tilburg University
Jean-David Fermanian	Professeur, ENSAE Paris

Institut Camille Jordan
UMR 5208 du CNRS

Modélisation de la dépendance entre pré-extrêmes

Simon Chatelain

Thèse de doctorat

McGill

Université Claude Bernard Lyon 1

À Yoye,

Remerciements

Les premières personnes que j'aimerais remercier sont bien sûr mes deux directrices de thèse pour m'avoir guidé tout au long de ma thèse. Merci Anne-Laure Fougères pour ton soutien continu, ta bienveillance, et ta vue d'ensemble qui m'a maintes fois sorti du bourbier. Merci Johanna G. Nešlehová de m'avoir initié si tôt à la recherche, de m'avoir encouragé à poursuivre mes études et d'avoir été si redoutable quand les plus grands obstacles se sont dressés face à nous. Je vous suis particulièrement redevable, à toutes les deux, pour la patience dont vous avez fait preuve pendant ces années de collaboration.

Ensuite, je voudrais remercier chaleureusement chacun des membres du jury pour l'intérêt que vous avez porté à mon travail. Je suis très reconnaissant envers Jean-David Fermanian et John Einmahl pour avoir accepté d'être les rapporteurs de ma thèse, votre présence et votre lecture détaillée du manuscrit m'honorent. J'en profite pour remercier aussi Frank Wagner et Philippe Naveau pour avoir suivi mon parcours et autorisé mes nombreuses réinscriptions.

Organiser une cotutelle n'est pas une tâche facile. Raffaella Bruno, Russell Steele, Helen Aaron et Jason Stillman ont su m'aider dans mes démarches à McGill tandis que le personnel administratif et technique de l'ICJ, en particulier Aurélie Raymond, Naima Fantoussa, Laurent Azema et Vincent Farget, m'a si bien accueilli dans le laboratoire. Je suis très reconnaissant de cette bienveillance qui m'a accompagné d'un continent à l'autre.

Deux universités, deux bureaux et quatre années (pour arrondir) de thèse c'est beaucoup de chercheurs, enseignants, doctorants et co-bureaux avec qui j'ai eu la chance de discuter autour d'un café. La liste est bien trop longue et je suis sûr que vous vous reconnaîtrez parmi les moments de convivialité qui ont rendu cette thèse humaine et agréable. Un grand merci à Maxime T., Quentin S., Simon Z., Léo B., Sam P. et Nicholas B. pour avoir apporté un souffle amical dans ce monde de maths. Merci Maxime H., ton enthousiasme infaillible m'impressionne, j'étais ravi que tu tiennes ta promesse d'une n-ième pause café à Montréal.

La thèse, c'est aussi tout ce qui se passe en dehors. Là encore, la liste est longue, mais je me sens si redevable envers celles et ceux qui m'ont nourri pendant ces quatre années que je n'épargnerai personne. Merci Yasmin pour m'avoir accompagné tout au long de ces années et pour m'avoir sorti de tous mes creux. Maman, Papa et Blaise pour le soutien constant et la confiance que vous m'avez accordé. Merci aux tribus Chatelain et Larisse. La famille Jhabvala aussi, pour votre accueil chaleureux. Merci à la coloc' de l'impasse, je pense particulièrement à nos belles échappées. Vincent pour nos fous-rires, Thomas pour ta gentillesse et Louis pour tes tartes. Chris. Elsa, je te trouve bien courageuse. Merci à ma famille Montréalaise avec qui j'ai passé deux années épanouies. Fede mon psy, mon messie. Ma douce Julie. Arnaud pour avoir pété mon lit. Les grimpeurs de Shakti, les ontariens et la maison Henri-Julien. Alyssa, Max et Leah pour le réconfort hebdomadaire. Gaëlle et ton énergie nucléaire. Merci surtout au parti Drolétaire. Merci Gabriel pour ce clavier. Léo pour la dose quotidienne de sérénité et les aventures aquatiques. Lison pour ta présence rassurante. Roman, drôle d'aide-soignant. Obrigado Day, pour toutes ces belles découvertes brésiliennes. Les soirées Bloomfield, la compagnie du théâtre circonspect. Alex et la disco Durocher. Clarence, merci pour ta complicité et ta passion de vie contagieuse. Merci, d'autant plus, à toi et à Léo pour avoir relu mon introduction et corrigé toutes mes fautes de farppe. Merci ici et merci là-bas. C'était bien fin.

Résumé

Le comportement extrême joint entre variables aléatoires revêt un intérêt particulier dans de nombreuses applications allant des sciences de l'environnement à la gestion du risque. Par exemple, ce comportement joue un rôle central dans l'évaluation des risques de catastrophes naturelles. Une erreur de spécification de la dépendance entre des variables aléatoires peut engendrer une sous-estimation dangereuse du risque, en particulier au niveau extrême. Le premier objectif de cette thèse est de développer des techniques d'inférence pour les copules Archimax. Ces modèles de dépendance peuvent capturer tout type de dépendance asymptotique entre les extrêmes et, de manière simultanée, modéliser les risques joints au niveau moyen. Une copule Archimax $C_{\psi,\ell}$ est caractérisée par ses deux paramètres fonctionnels, la fonction de dépendance caudale stable ℓ et le générateur Archimédien ψ qui agit comme une distorsion affectant le régime de dépendance extrême. Des conditions sont dérivées afin que ψ et ℓ soient identifiables, de sorte qu'une approche d'inférence semi-paramétrique puisse être développée. Deux estimateurs non paramétriques de ℓ et un estimateur de ψ basé sur les moments, supposant que ce dernier appartient à une famille paramétrique, sont avancés. Le comportement asymptotique de ces estimateurs est ensuite établi sous des hypothèses de régularité non restrictives et la performance en échantillon fini est évaluée par le biais d'une étude de simulation. Une construction hiérarchique (ou en "clusters") généralisant les copules Archimax est proposée afin d'apporter davantage de flexibilité, la rendant plus adaptée aux applications pratiques. Le comportement extrême de ce nouveau modèle de dépendance est étudié, ce qui engendre un nouvelle manière de construire des fonctions de dépendance caudale stable. La copule Archimax est ensuite utilisée pour analyser les maxima mensuels de précipitations observées à trois stations météorologiques en Bretagne. Le modèle semble très bien ajusté aux données, aussi bien aux précipitations faibles qu'aux fortes. L'estimateur non paramétrique de ℓ révèle une dépendance extrême asymétrique entre les stations, ce qui reflète le déplacement des orages dans la région. Une application du modèle Archimax hiérarchique à un jeu de données de précipitations contenant 155 stations est ensuite présentée, dans laquelle des groupes de stations

asymptotiquement dépendantes sont déterminés via un algorithme de “clustering” spécifiquement adapté au modèle. Enfin, de possibles méthodes pour modéliser la dépendance inter-cluster sont évoquées.

Abstract

In various applications in environmental sciences, finance, insurance or risk management, joint extremal behavior between random variables is of particular interest. For example, this plays a central role in assessing risks of natural disasters. Misspecification of the dependence between random variables can lead to substantial underestimation of risk, especially at extreme levels. This thesis develops inference techniques for Archimax copulas. These copula models can account for any type of asymptotic dependence between extremes and at the same time capture joint risks at medium levels. An Archimax copula $C_{\psi,\ell}$ is characterized by two functional parameters, the stable tail dependence function ℓ, and the Archimedean generator ψ which acts as a distortion of the extreme-value dependence model. Conditions under which ψ and ℓ are identifiable are derived so that a semiparametric approach for inference can be developed. Two nonparametric estimators of ℓ and a moment-based estimator of ψ, which assumes that the latter belongs to a parametric family, are proposed. The asymptotic behavior of the estimators is then established under broad regularity conditions; performance in small samples is assessed through a comprehensive simulation study. In the second part of the thesis, Archimax copulas are generalized to a clustered constructions in order to bring in more flexibility, which is needed in practical applications. The extremal behavior of this new dependence model is derived. Finally, the methodology proposed herein is illustrated on precipitation data. First, a trivariate Archimax copula is used to analyze monthly rainfall maxima at three stations in French Brittany. The model is seen to fit the data very well, both in the lower and in the upper tail. The nonparametric estimator of ℓ reveals asymmetric extremal dependence between the stations, which reflects heavy precipitation patterns in the area. An application of the clustered Archimax model to a precipitation dataset containing 155 stations is then presented, where groups of asymptotically dependent stations are determined via a specifically tailored clustering algorithm. Finally, possible ways to model inter cluster dependence are discussed.

Contributions to original knowledge

Chapter 3

First, this chapter establishes conditions under which Archimax copulas are identifiable, which is original scholarship. Regularity conditions on the two functional parameters are also proposed and shown to ensure smoothness of the Archimax copula. This was not addressed previously in the literaturee.

Chapter 4

This chapter contains a two new nonparametric estimators for the stable tail dependence function of Archimax copulas. These are novel estimators and can be seen as generalizations of the Pickands and CFG estimators for extreme-value copulas. The proof of weak convergence of these estimators is also original scholarship, as is the finite sample simulations study at the end of the chapter.

Chapter 5

The moment-based parametric estimator for the Archimedean generator of an Archimax copula proposed in this chapter is new. The asymptotic behavior which is established for the nonparametric estimator of the stable tail dependence function given that the Archimedean generator is unknown is original scholarship as well.

Chapter 6

The clustered Archimax copula proposed in this chapter is new. Its extremal behavior that is elicited and proved is also new.

Chapter 7

The application of Archimax copulas to the two precipitation datasets is original work. The pilot study contained in this thesis is also original. The clustering algorithm used to find groups of stations is the well known PAM algorithm with a distance which is original and tailored to the clustered Archimax copula. Proposition 7.1 is also original scholarship.

Contributions of authors

I played an integral part in the work done in Chapters 3, 4, 5 and Section 7.1. Namely, I established conditions on the two functional parameters of Archimax copulas for them to be identifiable, as well as conditions to ensure their regularity. I devised the nonparametric estimator for the stable tail dependence function, established their convergence and conducted the extensive simulation study found in Chapter 4. I also worked on the moment-based estimator presented in Chapter 5 and the proofs of the extensions of the weak convergence results therein. The application to precipitation data in Section 7.1, as well as the pilot study, is my work too. All these results were obtained under the supervision of my two Ph.D. supervisors Johanna G. Nešlehová and Anne-Laure Fougères. The corresponding paper, "Inference for Archimax copulas" was written by the three of us.

Chapter 6 consists of work that I did to extend the Archimax family of copulas to a more general clustered Archimax copula. The proposed model is original and I derived its extremal behavior, including a conjectured extension, as well as the corollaries and examples presented in the chapter. This work was supervised by Johanna G. Nešlehová and Anne-Laure Fougères; we have an upcoming paper about this model.

Section 7.2 is based on a working paper in preparation co-authored with Samuel Perreault of Université Laval. I am responsible for the majority of the work presented in this chapter. Namely, I selected the dataset, how it would be used, and which type of clustering algorithm to use, and worked out how the EM algorithm can be applied to estimate the dependence parameter of the distortions. The distance used in 7.2.1 is a result of discussion between Samuel and myself. The computational implementation of the EM algorithm, which is not a part of this thesis, is currently being worked on by both of us. None of this work will feature in Samuel's Ph.D. thesis.

Contents

Chapter 1

Introduction

Extreme environmental events such as floods, heat waves and cold spells can have catastrophic effects on the natural world and human society in the form of loss of infrastructure, capital and life. High precipitation in the province of Quebec has caused devastating floods when coupled with snow melt during the spring season. In the Cévennes region of southern France, moisture accumulated on the Mediterranean Sea over the summer is blown onto the region and trapped on the mountainside. This leads to stationary extreme precipitation events during the Fall season known as "Orages cévenols", which in turn cause destructive floods. It is important for public safety to be able to predict the risks of environmental disasters in order to establish preventive measures.

Extreme value theory had been a growing area of research since the first half of the 20th century. Developments in the area were first motivated by environmental applications. The gargantuan Delta Works designed to protect the Netherlands from storm surges are an excellent example of this. A large amount of infrastructure such as storm surge barriers, dams and levees were planned and built in the aftermath of the North Sea Flood of 1953 which devastated the Netherlands, Belgium and England. Nearly 2000 deaths were reported in the Netherlands alone. The most ambitious part of the Delta Works is a 9 kilometer long dam called the Oosterscheldekering which was designed to guarantee the safety of the population of Rotterdam for an event with a return period of 10,000 years (4,000 for Zeeland). Roughly speaking, an event with a return period of 10,000 years is defined as an event exceeding a certain threshold in a year with probability 1/10,000. Such extreme events are difficult to model due to the fact that they are also, by definition, extremely rare. While traditional statistical problems require modeling the center (or bulk) of the observed data, we find ourselves needing to model the tails of the distribution where information is scarce. This issue is especially prevalent in the field of environmental sciences where measurements of high quality rarely go back very far in time. How can one determine a 10,000 year return level with, say, 40 years of observations? To alleviate this fundamental issue, extreme value theory draws on more mathematical tools than other areas of statistics: stability properties and asymptotic behaviors are sought in order to extrapolate and infer on events which often fall outside the observed range of the data.

In the case of the North Sea Flood of 1953, on the night of the 31st of January, the sea level rose more than 5.6 meters above its average value in several locations. At first glance this event can be seen as a univariate statistical problem where the variable of interest is simply the sea level anomaly at a given location. However, a storm surge is caused by a combination of wind, high tide and low sea surface pressure. Moreover, the fact that we are often interested in quantifying risk at not one but multiple locations makes the problem all the more multi-dimensional. Indeed, understanding the behavior of each of the variables individually such as wind speed and pressure won't paint the whole picture. A crucial part of analyzing risk in environmental applications is to quantify the dependence between the variables of interest. The purpose of this thesis is to contribute to the array of tools available to model multivariate extremes, specifically focusing on modeling dependence.

The problem of sparsity of extremes is amplified in the context of dependence modeling. While the expression "curse of dimensionality" refers to various issues surrounding high dimensions in statistical learning, it is also partivularly relevant to our setting. To infer the dependence structure between several random variables, say d of them, one needs to have a sufficiently large sample in order to adequately fill the d-dimensional observation space. With extreme values being so few in occurrence, it could seem like an impossible feat, especially when dealing with hydro-meteorological applications. One popular approach to this issue is to impose a parametric model on the data which also greatly facilitates inference since well studied likelihood based methods can be applied with good quantification of uncertainty. Another solution is dimension reduction, which has recently garnered interest in the field of extreme value analysis. The idea explored in this thesis is instead to lower the barrier to what is considered an extreme in order to retain a larger portion of the dataset at hand. Traditionally, data points are selected to be extreme enough to apply models that are asymptotically justified, i.e. Generalized Extreme Value or Generalized Pareto univariate distributions tied together by so-called extreme-value copulas. Real datasets being finite in size, this is never verified but can be checked to be a reasonable modeling assumption to make.

Here, the asymptotic modeling assumption is relaxed. The terms subasymptotic (or pre-asymptotic) can have different meanings, in this thesis the intended definition is that the data is not deemed "extreme enough" to use asymptotic models. Instead of studying the class of extreme-value (or max-stable) copulas, the more general Archimax family is considered. Archimax copulas have the advantage of being particularly flexible. Foremost, it is fully flexible in the extreme regime, meaning that any asymptotic dependence structure can be attained by a subclass of Archimax copulas. The size of the family allows to simultaneously model dependence at medium levels as well. In fact, other desirable properties such as asymmetry and lower tail dependence are also possible to capture. While Archimax copulas have been known for some time, lack of proper inference tools have left the family rarely used in practice. The first goal of this thesis is to develop inference techniques for this family and evaluate their performance through convergence results, simulation studies and applications. The second goal is to expand the class to a hierarchical construction, in order to allow for even more flexible modeling of clustered data. Indeed, while being able to capture asymptotic dependence is necessary, it can also be of interest to additionally allow for asymptotic independence. This is possible in the hierarchical Archimax model, where asymptotic dependence and independence is possible within and between clusters. Clustering in multivariate extremes finds its use not only in exploratory data analysis but can also be employed to pool data in a spatial setting between asymptotically dependent stations.

All preliminary notions needed to understand the original research presented in this thesis can be found in Chapter 2. Namely, dependence modeling via copulas is presented, along with the Archimedean, extreme-value and Archimax families. Concepts of weak convergence for empirical processes are also presented as they are used later in the thesis. Essential properties of the Archimax family of copulas, namely identifiability and smoothness, are elicited in Chapter 3. This chapter verifies that powerful theorems can be applied to justify the inference tools developed herein, and it is thus often referred to in statements of important results throughout. Chapter 4 develops a non-parametric estimator for one of the two functional parameters of the Archimax copula, namely the stable tail dependence function. While not directly applicable to a real dataset, essential results concerning the asymptotic behavior of the estimation techniques are proved here. Small sample performance is also assessed via an extensive simulation study, whose detailed results can be found in Appendix A. Chapter 4 serves as a stepping stone to Chapter 5, where full inference for Archimax copulas is developed. Indeed, a moment-based procedure is proposed to estimate the other functional parameter, the Archimedean generator. The nonparametric approach of the previous chapter thus completes the procedure, hence the title of Chapter 5, "Semiparametric inference for Archimax copulas". Convergence results which are involved extensions of those from Chapter 4 are also obtained. Chapter 6 presents a new hierarchical (or clustered) Archimax model which addresses some shortcomings of the simple Archimax model. This allows to broaden the applications, while offering interpretability and preserving the strengths of the Archimax copula. The behavior of the model at the extreme regime is studied and points toward a new way to build dependence structures for extremes. Applications to real datasets are gathered in Chapter 7. First, a trivariate precipitation dataset is studied to illustrate the methodology developed in Chapter 5. The Archimax approach to assessing joint risk is compared to other techniques and thanks to a pilot simulation study, it is shown to be advantageous in certain situations. The scope of the dataset is then dramatically broadened from three to over a hundred stations in France. In order to model the precipitation amounts over this large geographical area, the hierarchical model from Chapter 6 is applied thanks to a clustering algorithm tailored to it. Finally, Chapter 8 concludes this thesis with a discussion and possible directions for future work.

En Français

Les événements environnementaux extrêmes tels que les inondations et les vagues de chaleur ont des effets catastrophiques sur les milieux naturels ainsi que sur la société humaine en matière de perte d'infrastructure, de capital et de vie. Par exemple, des précipitations extrêmes au Québec causent des inondations dévastatrices lorsqu'elles sont combinées aux fontes des neiges printanières. Dans la région des Cévennes en France, l'humidité accumulée durant l'été à la surface de la mer Méditerranée est acheminée au dessus de la région par des vents venant du sud, provoquant ainsi des orages stationnaires. Ces orages, appelés "orages cévenols", sont connus pour leur conséquences destructrices. Il est donc important, pour des questions de sécurité publique, de pouvoir prédire les risques de catastrophes environnementales afin d'établir des mesures de prévention et de protection.

La théorie des valeurs extrêmes est un domaine de recherche qui connait une forte croissance depuis la première moitié du vingtième siècle. Ce développement fut principalement motivé par des applications environnementales : le gargantuesque projet Delta conçu pour protéger les Pays-Bas des inondations maritimes en est un parfait exemple. Il comprend de nombreuses infrastructures, notamment des barrages, des digues et des clôtures, planifiées et réalisées suite au raz-de-marée de 1953 en Mer du Nord. Cette année-là, le raz-de-marée causa la mort d'environ 2000 personnes. La construction la plus ambitieuse de ce projet est un barrage long de 9 km, appelé Oosterscheldekering. Il a été pensé pour protéger la population de Rotterdam contre un événement dont la période de retour est de 10 000 années (4000 années pour la population de la Zélande). De manière simplifiée, on définit un événement avec une période de retour de 10 000 années par le seuil dépassé, en une année donnée, avec une probabilité de 1/10 000. De tels événements sont difficiles à modéliser statistiquement dans la mesure où ils sont, par définition, extrêmement rares. Si les problèmes statistiques traditionnels requièrent souvent de modéliser le centre des données observées, ici le besoin est plutôt celui de modéliser les queues des distributions, là où l'information est très peu abondante. Ce manque d'information est d'autant plus présent dans les applications environnementales où les séries de mesures de quantités physiques, telles que des débits d'eau, sont souvent courtes ou de qualité médiocre. Comment déterminer un événement avec une période de retour de 10 000 années avec seulement 40 années d'observations ? Pour pallier cette difficulté, la théorie des valeurs extrêmes emprunte de nombreux outils mathématiques en comparaison à d'autres domaines de la statistique. En effet, on recherche des propriétés de stabilité et des comportements asymptotiques afin de pouvoir extrapoler et inférer des valeurs qui sortent souvent du champ des données observé.

Si on reprend l'exemple du raz-de-marée de 1953, dans la nuit du 31 janvier, le niveau de la mer s'est élevé de plus de 5,6 mètres au-dessus du niveau moyen, et ceci à plusieurs endroits le long de la côte Néerlandaise. À première vue, on pourrait croire qu'il s'agit d'un problème statistique univarié, où la variable d'intérêt est simplement l'anomalie du niveau de la mer en un lieu donné. Or, les raz-de-marée sont causés par une combinaison de vent, de haute marée et de basse pression atmosphérique. Ajoutons à cela le fait que, la plupart du temps, il est nécéssaire d'évaluer le risque en plusieurs lieux différents, il est évident que le problème en est d'autant plus multidimensionnel. En effet, étudier chaque variable individuellement ne permettra pas de dresser un portrait complet du phénomène, c'est pourquoi lors de l'analyse du risque dans les sciences environnementales, il est crucial de quantifier la dépendance entre les variables d'intérêt. L'objectif de cette thèse est de contribuer à l'éventail des outils permettant de modéliser les extrêmes multivariés, et particulièrement la dépendance entre ceux-ci.

La sparsité des valeurs extrêmes est exacerbée dans le contexte multivarié, de fait, l'expression courante du “fléau de la dimension” est pertinente ici. Afin d'inférer la structure de dépendance entre plusieurs variables aléatoires, disons d d'entre elles, il nous faut un échantillon de données suffisamment grand pour couvrir l'espace d'observation à d dimensions. Étant donnée la rareté inhérente aux événements extrêmes, ceci peut sembler être une cause perdue surtout dans le domaine hydrométéorologique, qui, comme nous l'avons précisé plus tôt, est un domaine qui manque de données. Une approche courante est d'imposer un modèle paramétrique sur les valeurs extrêmes du jeu de données, ce qui facilite grandement l'inférence grâce à l'abondance de résultats déjà établis sur les méthodes d'ajustement par maximum de vraisemblance. Celles-ci permettent une bonne quantification de l'incertitude, qualité également présente dans les méthodes bayésiennes. Une autre approche assez populaire aujourd'hui consiste à effectuer une réduction de dimension. L'idée avancée par cette thèse est plutôt d'élargir la classe d'événements considérés comme étant extrêmes, afin de conserver une plus grande proportion des données disponibles. Traditionnellement, on sélectionne les observations suffisamment extrêmes pour ajuster des modèles asymptotiquement justifiés, tels que des lois de valeurs extrêmes généralisées, liées par des copules de valeurs extrêmes. Les jeux de données étant finis, ils ne peuvent jamais être parfaitement décrits par de tels modèles, bien qu'il existe des méthodes pour vérifier si leur utilisation est judicieuse.

Dans cette thèse, le régime asymptotique n'est pas imposé. L'expression subasymptotique (ou pré-asymptotique) a différentes significations, ici, elle indique le fait que les données ne sont pas suffisamment extrêmes pour employer des modèles asymptotiques. Nous nous pencherons sur une famille de copules, appelée Archimax, qui généralise les copules de valeurs extrêmes communément utilisées dans ce domaine. La classe Archimax a l'avantage d'être très flexible. D'une part, cette flexibilité est présente dans son comportement extrême, puisque n'importe quelle structure de dépendance asymptotique peut être atteinte par une sous-classe de copules Archimax. D'autre part, la grandeur de cette famille permet de modéliser de manière simultanée la dépendance à plusieurs niveaux. De plus, d'autres propriétés désirables, comme l'asymétrie et la présence de dépendance caudale inférieure, peuvent également être capturées. Bien que cette famille soit connue depuis un certain temps, le manque d'outils d'inférence a limité son utilisation dans des contextes applicatifs. Le premier objectif de cette thèse est donc de développer des techniques permettant d'ajuster des lois Archimax et d'en étudier les propriétés à travers des résultats de convergence, des simulations et des applications à des données réelles. Le deuxième objectif est d'élargir cette classe de distributions, grâce à une construction hiérarchique, afin d'apporter plus de flexibilité. Effectivement, bien que la dépendance asymptotique soit un régime important à modéliser, il est aussi intéressant de capturer l'indépendance asymptotique. Ceci est rendu possible grâce au modèle hiérarchique proposé par cette thèse. Plus précisément, le modèle permet de lier plusieurs clusters de variables, avec suffisamment de flexibilité pour permettre à la fois de la dépendance et de l'indépendance inter et intra-cluster. Ce "clustering" est utile dans un contexte d'analyse exploratoire des données mais peut également être utilisé plus largement, notamment pour mettre en commun des variables ayant un comportement extrême semblable.

Toutes les notions préliminaires nécessaires à la compréhension de cette thèse sont présentées dans le Chapitre 2. Il contient une section sur la modélisation de la dépendance, présentant ainsi les trois familles de copules importantes pour nous: les Archimédiennes, celles de valeurs extrêmes et les Archimax. En deuxième partie, le Chapitre 2 développe le concept de convergence faible pour les processus empiriques, nécessaires aux résultats théoriques des chapitres suivants. Des propriétés essentielles de régularité sont étudiées dans le Chapitre 3. Celui-ci permet de vérifier que certains théorèmes fins peuvent être appliqués aux méthodes développées et ainsi apporter une justification théorique. Le Chapitre 4 propose une méthode d'estimation non paramétrique pour l'un des deux paramètres fonctionnels de la copule Archimax, la fonction de dépendance caudale stable. Bien qu'elle ne puisse pas être directement appliquée à un jeu de données, des résultats essentiels concernant son comportement asymptotique sont prouvés et sa performance en échantillon fini est également étudiée et détaillée en Annexe A. Le Chapitre 4 pose les bases d'une inférence complète pour les copules Archimax que nous développerons dans le Chapitre 5. En effet, on y trouve une estimation par moments de l'autre paramètre fonctionnel, le générateur Archimédien. La méthode non paramétrique du Chapitre 4 va ainsi compléter la procédure d'ajustement, expliquant le titre du Chapitre 5, "Inférence semi-paramétrique pour copules Archimax". Nous obtiendrons des résultats de convergence, versions généralisées des résultats du Chapitre 4. Ensuite, le Chapitre 6 propose un nouveau modèle Archimax hiérarchique pour combler certaines lacunes du modèle Archimax simple. Il permet d'élargir les possibilités d'application en offrant une interprétabilité intéressante tout en conservant les atouts des copules Archimax. Nous y étudierons le comportement extrême du modèle et suggérerons une nouvelle méthode pour construire des structures de dépendance de valeurs extrêmes. Les applications à des jeux de données réelles se trouvent dans le Chapitre 7. En première partie, un jeu de données de précipitations trivarié est utilisé pour illustrer la méthodologie développée dans le Chapitre 5. Ensuite, la modélisation par copule Archimax de risques extrêmes est comparée à d'autres techniques courantes grâce à une étude de simulation qui souligne ses avantages. En deuxième partie, le jeu de données est élargi à plus de cent cinquante stations météorologiques en France. Le modèle hiérarchique du Chapitre 6 est convoqué, notamment via un algorithme de clustering adapté, afin de modéliser les précipitations sur un territoire si grand. Le Chapitre 8 conclut cette thèse par une discussion et des perspectives de recherche futures.

Chapter 2

Background

This chapter contains all necessary background information needed to read the chapters that follow. Section 2.1 treats the subject of copulas, most importantly defining the Archimax family in Section 2.1.4 which are studied in depth in this thesis. Section 2.2 defines the notions relating to weak convergence needed to validate non-parametric approaches to estimate copulas.

In what follows, vectors in $\mathbb{R}^d$ are denoted by boldface letters, viz. $\boldsymbol{x} = (x_1, \ldots, x_d)$. Binary operations such as $\boldsymbol{x} + \boldsymbol{y}$ or $a \cdot \boldsymbol{x}$, $\boldsymbol{x}^a$ are understood as component-wise operations. In particular, for any function $f : \mathbb{R} \to \mathbb{R}$ and $\boldsymbol{x} \in \mathbb{R}^d$, $f(\boldsymbol{x})$ denotes the vector $(f(x_1), \ldots, f(x_d))$. Furthermore, $\|\cdot\|$ stands for the ℓ_1-norm, viz. $\|\boldsymbol{x}\| = x_1 + \cdots + x_d$. For any $x, y \in \mathbb{R}$, let $x \wedge y = \min(x, y)$ and $x \vee y = \max(x, y)$. Finally, $\mathbb{R}^d_+$ is the positive orthant $[0, \infty)^d$ and for any $x \in \mathbb{R}$, x_+ denotes the positive part of x.

2.1 Copulas

A copula is simply a d-dimensional distribution function on the unit hypercube with uniform margins. A formal definition is given below.

Definition 2.1. *A d-dimensional copula is a function $C : [0,1]^d \to [0,1]$ satisfying*

(i) $C(u_1, \ldots, u_d) = 0$ *whenever* $u_j = 0$ *for at least one* $j \in \{1, \ldots, d\}$.

(ii) $C(u_1, \ldots, u_d) = u_j$ *if* $u_i = 1$ *for all* $i \in \{1, \ldots, d\}$ *and* $i \neq j$.

(iii) C *is d-nondecreasing on* $[0,1]^d$. *That is, for each hyperrectangle* $R = \prod_{j=1}^d [a_j, b_j] \subset [0,1]^d$, *the C-volume of R is nonnegative, i.e.*

$$\int_R dC(\boldsymbol{u}) = \sum_{i_1=1}^{2} \cdots \sum_{i_d=1}^{2} (-1)^{i_1+\ldots+i_d} C(u_{1i_1}, \ldots, u_{di_d}) \geq 0 \ ,$$

where for $j \in \{1, \ldots, d\}$, $u_{j1} = a_j$ *and* $u_{j2} = b_j$.

Copulas arose when probabilists were interested in the properties of multivariate distributions with given marginal distributions. Specifically, given d univariate distributions

$F_1, \ldots, F_d$, how can a d-dimensional distribution F be constructed so that the margins are precisely $F_1, \ldots, F_d$? Standardization of the marginals to a common distribution helps in isolating the underlying dependence structure. In the case of a continuous real random vector $(X_1, \ldots, X_d)$, applying the probability integral transforms component-wise, viz. $(F_1(X_1), \ldots, F_d(X_d))$, yields a random vector whose distribution is supported on the unit hypercube $[0,1]^d$ and has uniform margins. In the following theorem due to Sklar (1959), the link between F and the marginals $F_1, \ldots, F_d$ is established via copulas. The result below is stated for the case of continuous marginals, since extensions to discontinuous margins are not needed in this thesis. Such extensions can be found, for example, in Nelsen (2006) and Genest and Nešlehová (2007).

Theorem 2.1. *Let $d \in \mathbb{N}$, $d \geq 2$.*

- *Let F be a distribution function on $\mathbb{R}^d$ with continuous margins $F_1, \ldots, F_d$ and $\boldsymbol{X} = (X_1, \ldots, X_d) \sim F$. Then there exists a unique distribution function C on $[0,1]^d$ with uniform margins, named the copula of $\boldsymbol{X}$, such that, for all $\boldsymbol{x} = (x_1, \ldots, x_d) \in \mathbb{R}^d$,*

$$F(\boldsymbol{x}) = C(F_1(x_1), \ldots, F_d(x_d)) ,$$

  and C is defined for all $\boldsymbol{u} = (u_1, \ldots, u_d) \in [0,1]^d$ by

$$C(\boldsymbol{u}) = F(F_1^{-1}(u_1), \ldots, F_d^{-1}(u_d)) ,$$

  where for $j \in \{0, \ldots, d\}$, $F_j^{-1}(u_j) = \inf\{x_j \in \mathbb{R} : F_j(x_j) \geq u_j\}$ for $u_j \in [0,1]$.

- *Conversely, if $F_1, \ldots, F_d$ are distribution functions on $\mathbb{R}$, and C is a copula, then F as given for all $\boldsymbol{x} = (x_1, \ldots, x_d) \in \mathbb{R}^d$ by*

$$F(\boldsymbol{x}) = C(F_1(x_1), \ldots, F_d(x_d))$$

  is a joint distribution on $\mathbb{R}^d$ with copula C and marginal distributions $F_1, \ldots, F_d$.

The implications of the above theorem for dependence modeling are important. Indeed, it effectively separates marginal distributions from the underlying dependence structure characterized by the copula. This means that in practice, marginal effects can be modeled separately (usually before) modeling the dependence between them. This also means that given a set of marginal distributions, a variety of joint distributions can be created by tying them together with copulas. There is a vast amount of literature focusing on the use of copulas for dependence modeling in multivariate statistical problems. One can refer to the comprehensive monographs by Joe (2014) and Nelsen (2006). Copulas have been applied in many fields ranging such as hydrology (see Salvadori et al. (2007)), risk management (see McNeil et al. (2005)) and finance (see Mai and Scherer (2014) or Cherubini et al. (2004) for example).

An analogous theorem links multivariate survival functions to marginal survival functions via survival copulas. Survival copulas, denoted $\bar{C}$, are also copulas and are often employed in this thesis. One can refer to Chapter 2.6 in Nelsen (2006) for an overview in the bivariate case. The following result is also stated in the special case of continuity.

Theorem 2.2. *Let $d \in \mathbb{N}$, $d \geq 2$.*

- *Let $\bar{F}$ be a survival function on $\mathbb{R}^d$ with continuous marginals $\bar{F}_1, \ldots, \bar{F}_d$ and $\boldsymbol{X} = (X_1, \ldots, X_d) \sim F$. Then there exists a copula $\bar{C}$ on $[0,1]^d$ with uniform margins, named the survival copula of $\boldsymbol{X}$, such that, for all $\boldsymbol{x} = (x_1, \ldots, x_d) \in \mathbb{R}^d$,*

$$\bar{F}(\boldsymbol{x}) = \bar{C}(\bar{F}_1(x_1), \ldots, \bar{F}_d(x_d)) \ ,$$

and $\bar{C}$ is defined for all $\boldsymbol{u} = (u_1, \ldots, u_d) \in [0,1]^d$ by

$$\bar{C}(\boldsymbol{u}) = \bar{F}(F_1^{-1}(1-u_1), \ldots, F_d^{-1}(1-u_d)) \ .$$

- *Conversely, if $\bar{F}_1, \ldots, \bar{F}_d$ are continuous survival functions on $\mathbb{R}$, and $\bar{C}$ is a copula, then $\bar{F}$ as given for all $\boldsymbol{x} = (x_1, \ldots, x_d) \in \mathbb{R}^d$ by*

$$\bar{F}(\boldsymbol{x}) = \bar{C}(\bar{F}_1(x_1), \ldots, \bar{F}_d(x_d))$$

is a joint survival function on $\mathbb{R}^d$ with survival copula $\bar{C}$ and margins $\bar{F}_1, \ldots, \bar{F}_d$.

Suppose that $(X_1, \ldots, X_d)$ is a random vector with continuous margins $F_1, \ldots, F_d$, copula C and survival copula $\bar{C}$. Let $\tilde{C}(u_1, \ldots, u_d) = \Pr(F_1(X_1) > u_1, \ldots, F_d(X_d) > u_d)$. The survival copula is related to the copula through the following expression. For all $\boldsymbol{u} \in [0,1]^d$,

$$\bar{C}(\boldsymbol{u}) = \tilde{C}(1-u_1, \ldots, 1-u_d) \ ,$$

where $\tilde{C}$ can be written in terms of C viz.

$$\tilde{C}(\boldsymbol{u}) = \sum_{\iota_1, \ldots, \iota_d \in \{0,1\}} (-1)^{\iota_1 + \ldots + \iota_d} C(u_1 \vee \iota_1, \ldots, u_d \vee \iota_d) \ .$$

Conversely,

$$C(\boldsymbol{u}) = \sum_{\iota_1, \ldots, \iota_d \in \{0,1\}} (-1)^{\iota_1 + \ldots + \iota_d} \bar{C}(1 - u_1\iota_1, \ldots, 1 - u_d\iota_d) \ . \tag{2.1}$$

The following properties concerning copulas are helpful and used throughout this thesis. Let C be any d-dimensional copula of a random vector $(X_1, \ldots, X_d)$. Then,

(A) If $X_1, \ldots, X_d$ are continuous, then

$$X_1, \ldots, X_d \text{ are independent} \iff C(\boldsymbol{u}) = C_\Pi(\boldsymbol{u}) = u_1 \ldots u_d \ .$$

(B) (Fréchet-Hoeffding bounds) For all $\boldsymbol{u} \in [0,1]^d$,

$$\max\{1 - d + \sum_{j=1}^{d} u_j, 0\} = W(\boldsymbol{u}) \leq C(\boldsymbol{u}) \leq C_M(\boldsymbol{u}) = \min\{u_1, \ldots, u_d\} .$$

(C) C is Lipschitz continuous with respect to the ℓ_1 norm. That is, for $\boldsymbol{u}, \boldsymbol{v} \in [0,1]^d$,

$$|C(\boldsymbol{u}) - C(\boldsymbol{v})| \leq ||\boldsymbol{u} - \boldsymbol{v}||_1 = \sum_{j=1}^{d} |u_j - v_j| .$$

(D) Let $j \in \{1, \ldots, d\}$. Then the partial derivative $\dot{C}_j(\boldsymbol{u}) = \partial C(\boldsymbol{u})/\partial u_j$ exists for all $u_{j'} \in [0,1]$ and almost all $u_j \in [0,1]$, $j' \neq j$. Moreover, due to Lipschitz continuity, $0 \leq \dot{C}_j \leq 1$ wherever it exists.

These properties are proved, for example, in the monograph by Nelsen (2006).

Remark 2.1. *Note that the upper bound C_M in (B) above is a bona fide copula while the lower bound W is not for $d \geq 3$. In the case $d = 2$, W corresponds to perfect negative dependence, a concept which is not generalizable to higher dimensions. It is however a pointwise sharp bound. See Theorems 3.3 and 3.9 in Joe (2014), or Theorems 2.10.12 and 2.10.13 in Nelsen (2006).*

2.1.1 Measures of dependence

While copulas paint the whole picture regarding the dependence between several random variables, it is often of interest to report summarizing measures of dependence. Such dependence concepts are important to acquire an intuition about joint behavior of random variables and help communicate results of statistical analysis. In the following, we define the dependence measures used in this thesis. While generalizations to higher dimensions exist, they are best understood in the bivariate setting. Examples will be given in the subsequent sections regarding specific copula families.

Definition 2.2 (Rank correlation)**.** *Let X_1, X_2 be random variables with joint distribution F and marginal distribution functions F_1 and F_2. Spearman's rank correlation is given by*

$$\rho_S(X_1, X_2) = \rho(F_1(X_1), F_2(X_2)) ,$$

where ρ is the well-known Pearson's linear correlation. Let (X_1, X_2) and (X_1', X_2') be two independent realizations from F. Then Kendall's rank correlation (also called the coefficient of agreement, see Kendall and Babington Smith (1940)) is defined as

$$\tau(X_1, X_2) = \Pr[(X_1' - X_1)(X_2' - X_2) > 0] - \Pr[(X_1' - X_1)(X_2' - X_2) < 0] .$$

These two concepts of correlation avoid many pitfalls of the traditionally used linear correlation (see the cautionary article by Embrechts et al. (2002) for more details). Most relevant to this thesis, these measures do not depend on the margins, hence depending only on the underlying copula C. If the margins are continuous, they take following integral forms.

$$\begin{aligned}\rho_S(X_1, X_2) =& 12 \int_0^1 \int_0^1 \{C(u_1, u_2) - u_2 u_2\} du_1 du_2 \; , \\ \tau(X_1, X_2) =& 4 \int_0^1 \int_0^1 C(u_1, u_2) du_1 du_2 - 1 \; .\end{aligned}$$

This thesis being concerned with modeling at extreme levels, a measure of dependence which focuses on the tails of joint distributions is of interest. In the following, tail dependence coefficients are defined in the monograph by Joe (2014).

Definition 2.3. *Let X_1, X_2 be random variables with distributions F_1 and F_2. The coefficients of upper and lower tail dependence are*

$$\lambda_U = \lim_{q \uparrow 1} \Pr(F_2(X_2) > q | F_1(X_1) > q) \; , \tag{2.2}$$

$$\lambda_L = \lim_{q \downarrow 0} \Pr(F_2(X_2) < q | F_1(X_1) < q) \; . \tag{2.3}$$

provided the limits $\lambda_L, \lambda_U \in [0,1]$ exist. In the case of continuous margins, then noting that there is a unique copula C such that $(F_1(X_1), F_2(X_2)) = (U_1, U_2) \sim C$, $\lambda_U = 2 - \lim_{q\uparrow 1}\{1 - C(q,q)\}/(1-q)$ and $\lambda_L = \lim_{q\downarrow 0} C(q,q)/q$.

The pair (X_1, X_2) is said to be asymptotically dependent if $\lambda_U > 0$ and asymptotically independent if $\lambda_U = 0$. Since the case of asymptotic independence is reduced to only one point of the unit interval, a coefficient which allows to discriminate within this class of bivariate distributions is needed. Initially proposed by Ledford and Tawn (1996), residual tail dependence coefficients are introduced.

Definition 2.4. *Let X_1, X_2 be continuous random variables with distributions F_1 and F_2 and copula C. The residual upper and lower tail dependence indices*

$$\eta_U = \lim_{q \uparrow 1} \frac{\log(1-q)}{\log \tilde{C}(q,q)} \; , \qquad \eta_L = \lim_{q \downarrow 0} \frac{\log(q)}{\log C(q,q)} \; , \tag{2.4}$$

where $\tilde{C}(u_1, u_2) = 1 - u_1 - u_2 + C(u_1, u_2)$.

Here, $\eta_U, \eta_L \in [0,1]$, with 1 representing asymptotic dependence. Within asymptotic independence, Ledford and Tawn (1996) identify three types of dependence depending on where η falls within the unit interval.

Remark 2.2. *The two previously defined measures of tail dependence are, by definition, asymptotic. Functions that capture the penultimate tail behavior are also used, as is the case in Chapter 7.*

2.1.2 The Archimedean family

Archimedean copulas are a convenient and broadly studied class of copulas with many applications in areas such as finance and insurance. They are generated by a particular class of functions called Archimedean generators.

Definition 2.5. *A non-increasing and continuous function $\psi : [0,\infty) \to [0,1]$ which satisfies $\psi(0) = 1$, $\lim_{x\to\infty} \psi(x) = 0$ and is strictly decreasing on $[0, x_\psi)$, where $x_\psi = \inf\{x : \psi(x) = 0\}$, is called an Archimedean generator. By convention, $\psi(\infty) = 0$. The inverse $\phi : [0,1] \mapsto [0,\infty]$ of an Archimedean generator is defined as the inverse of ψ on $(0,1]$ and by $\phi(0) = x_\psi$.*

Archimedean copulas take the following form, for $\boldsymbol{u} \in [0,1]^d$ and an Archimedean generator ψ,

$$C_\psi(u_1,\dots,u_d) = \psi\left\{\phi(u_1) + \cdots + \phi(u_d)\right\}. \tag{2.5}$$

However, for this to be a copula, the notion of d-monotonicity is needed.

Definition 2.6. *An Archimedean generator ψ is called k-monotone, $k \in \mathbb{N}$ and $k \geq 2$, if it is differentiable on $(0,\infty)$ up to the order $k-2$, the derivatives satisfy $(-1)^m\psi^{(m)}(x) \geq 0$ for all $x \in (0,\infty)$ and $m \in \{1,\dots,k-2\}$, and further if $(-1)^{k-2}\psi^{(k-2)}$ is non-increasing and convex on $(0,\infty)$.*

Note that 2-monotone simply means that ψ is convex, and that a d-monotone Archimedean generator is also k-monotone for all $k \leq d$. McNeil and Nešlehová (2009) show that a function of the form (2.5) is a copula if and only if the generator ψ is d-monotone. It is also known that for an Archimedean generator to generate a copula in any dimension, it must be completely monotone, that is $(-1)^m\psi^{(m)}(x) \geq 0$ for all $m \in \mathbb{N}$ (see Kimberling (1974)). As will be explained shortly, the following transform due to Williamson (1956) is used to produce Archimedean generators from nonnegative random variables.

Definition 2.7. *If R is a nonnegative random variable with distribution F_R satisfying $F_R(0) = 0$ and $d \geq 2$ is an integer, then the Williamson d-transform of F_R is a real function defined for $x \in \mathbb{R}_+$ by*

$$\mathfrak{W}_d F_R(x) = \int_x^\infty \left(1 - \frac{x}{r}\right)^{d-1} dF_R(r) = \begin{cases} \mathrm{E}\left(1 - \frac{x}{R}\right)_+^{d-1} & \text{if} \quad x > 0 \\ 1 - F_R(0) & \text{if} \quad x = 0 \end{cases}.$$

As shown in Proposition 3.1 by McNeil and Nešlehová (2009), the distribution of a nonnegative random variable is uniquely given by its Williamson d-transform. Moreover, if $f = \mathfrak{W}_d F_R$, then for $x \in \mathbb{R}_+$, $F_R(x) = \mathfrak{W}_d^{-1} f(x)$ where

$$\mathfrak{W}_d^{-1} f(x) = 1 - \sum_{k=0}^{d-2} \frac{(-1)^k x^k f^{(k)}(x)}{k!} - \frac{(-1)^{(d-1)} x^{d-1} f_+^{(d-1)}(x)}{(d-1)!}. \tag{2.6}$$

Table 2.1: Archimedean generators and their dependence measures. D_1 denotes the Debye function (see Chapter 27 from Abramowitz and Stegun (1964)). $(\star)$ corresponds to the analytic form $1 - 4\sum_{k=1}^{\infty} 1/(k(\theta k+2)(\theta(k-1)+2))$. (†) indicates that $\eta_U = 1/2$ if $\theta = 1$, and (‡) that $\eta_L = 1/2$ if $\theta = 0$.

Family	$\psi_\theta(x)$	$\mathcal{O}$	τ	λ_U	λ_L	η_U	η_L
Clayton	$(1+\theta x)^{-1/\theta}$	$(0,\infty)$	$\theta/(\theta+2)$	0	$2^{-1/\theta}$	1/2	1(‡)
Frank	$-(1/\theta)\log\{1+e^{-x}(e^{-\theta}-1)\}$	$\mathbb{R}$	$1-4/\theta(1-D_1(\theta))$	0	0	1/2	1/2
Gumbel	$\exp(-x^{1/\theta})$	$[1,\infty)$	$\theta/(\theta+1)$	$2-2^{1/\theta}$	0	1(†)	$1-1/\theta$
Joe	$1-\{1-e^{-x}\}^{1/\theta}$	$[1,\infty)$	$(\star)$	$2-2^{1/\theta}$	0	1(†)	1/2

Another important notion in order to elicit the stochastic representation of Archimedean copulas is the class of ℓ_1-norm symmetric distributions. In the following, the unit simplex is defined as

$$\Delta_d = \left\{ \boldsymbol{s} \in \mathbb{R}_+^d : ||\boldsymbol{s}||_1 = 1 \right\} .$$

Definition 2.8. *A random vector $\boldsymbol{X}$ on $\mathbb{R}_+^d$ follows an ℓ_1-norm symmetric distribution if and only if there exists a nonnegative random variable R independent of $\boldsymbol{S}_d$ where $\boldsymbol{S}_d$ is a random vector uniformly distributed on the unit simplex so that $\boldsymbol{X}$ permits the stochastic representation*

$$\boldsymbol{X} \stackrel{d}{=} R\boldsymbol{S}_d .$$

All the elements needed for the stochastic representation being defined, Theorem 3.1 from McNeil and Nešlehová (2009) is reproduced below.

Theorem 2.3. *(i) Let $\boldsymbol{X}$ have a d-dimensional ℓ_1-norm symmetric distribution with radial distribution F_R satisfying $F_R(0) = 0$. Then $\boldsymbol{X}$ has an Archimedean survival copula with generator $\psi = \mathfrak{W}_d F_R$.*

(ii) Let $\boldsymbol{U}$ be distributed according to the d-dimensional Archimedean copula C_ψ with generator ψ (itself having the inverse ϕ). Then $(\phi(U_1), \ldots, \phi(U_d))$ has an ℓ_1-norm symmetric distribution with survival copula C_ψ and radial distribution F_R satisfying $F_R = \mathfrak{W}_d^{-1}\psi$.

This stochastic representation allows to create a variety of Archimedean copulas and sample from them. Table 2.1 presents a limited selection of generators and their dependence measures. For a larger variety, one can refer to Table 4.1 from Nelsen (2006) for example. The tail behavior of Archimedean copulas was extensively studied by Charpentier and Segers (2009), Table 2.1 only reports the measures of dependence presented before.

2.1.3 The Extreme-Value family

This section introduces the family of extreme-value copulas, which are crucial to the work presented in this thesis. However, before defining them, important results in univariate extreme value theory are given. Indeed, these results for univariate random variables are often called upon in the later chapters of this thesis. Comprehensive books on this subject include those from Resnick (1987), Coles (2001), Beirlant et al. (2004), Embrechts et al. (1997) and de Haan and Ferreira (2006).

A natural approach to statistical analysis of extremes, the so-called block-maxima approach, is to study the distribution of the maximum of n independent and identically distributed random variables $X_1, \ldots, X_n \sim F$. The variable of interest $M_n = \max\{X_1, \ldots, X_n\}$ is often taken over a block size motivated by the specific problem at hand, and large enough to warrant the use of an extreme distribution. For environmental applications, yearly or seasonal maxima are often considered. However, as $n \to \infty$ the distribution of M_n, which is equal to F^n, converges to a degenerate limit with point mass at the upper end-point x_F of the support of F, viz $x_f = \sup\{x \in \mathbb{R} : F(x) < 1\}$. It is therefore useful to find normalizing sequences $a_n > 0$ and $b_n \in \mathbb{R}$ such that for all x,

$$\lim_{n\to\infty} \Pr\left(\frac{M_n - b_n}{a_n} \le x\right) = G(x) \, ,$$

for some non-degenerate distribution G. If the above limit does exist, then F is said to be in the maximum domain of attraction of G, which is denoted $F \in \mathcal{M}(G)$ in this thesis. The possible forms G can take were determined by Fisher and Tippett (1928) and proved by Gnedenko (1943). Before stating the said theorem, recall the notion of regular variation.

Definition 2.9. *$f : \mathbb{R}_+ \mapsto \mathbb{R}_+$ is regularly varying with index $\alpha \in \mathbb{R}$ if and only if for all $t > 0$,*

$$f(xt)/f(x) \to t^\alpha$$

as $x \to \infty$, in notation $f \in \mathcal{R}_\alpha$.

Theorem 2.4 (Fisher-Tippett-Gnedenko Theorem)**.** *Let $X_1, \ldots, X_n$ be i.i.d. random variables with distribution F. Let $a_n > 0$ and $b_n \in \mathbb{R}$ be sequences such that $\lim_{n\to\infty} \Pr((M_n - b_n)/a_n \le x) = G(x)$ for a non-degenerate G and all continuity points x of G. Then, up to location and scale, for $\alpha > 0$, G is of one of the following three forms:*

(Fréchet) For $x \in \mathbb{R}$,

$$\Phi_\alpha(x) = \exp(-x^{-\alpha})\mathbf{1}(x \ge 0) \, ,$$

and $F \in \mathcal{M}(\Phi_\alpha)$ if and only if for all $t > 0$,

$$\lim_{x\to\infty} \frac{1 - F(tx)}{1 - F(x)} = t^{-\alpha} \, .$$

that is, if and only if $\bar{F} \in \mathcal{R}_{-\alpha}$.

(Gumbel) *For* $x \in \mathbb{R}$,

$$\Lambda(x) = \exp\{-\exp(-x)\} ,$$

and $F \in \mathcal{M}(\Lambda)$ *if and only if for some positive function* a, *for all* $t > 0$,

$$\lim_{x \to x_F} \frac{1 - F(x + ta(x))}{1 - F(x)} = e^{-t} ,$$

where x_F *is the upper end-point of the support of* F.

(Weibull) *For* $x \in \mathbb{R}$,

$$\Psi_\alpha(x) = \begin{cases} \exp\{-|x|^\alpha\} & \text{if} \quad x \leq 0 \\ 1 & \text{if} \quad x > 0 \end{cases} ,$$

and $F \in \mathcal{M}(\Psi_\alpha)$ *if and only if* $x_F < \infty$ *and for all* $t > 0$,

$$\lim_{x \to \infty} \frac{1 - F(x_F - \{tx\}^{-1})}{1 - F(x_F - \{x\}^{-1})} = t^{-\alpha} .$$

For the Gumbel domain of attraction, the function a, called an auxiliary function, is not unique. It can be chosen to be $\int_x^{x_\star} \bar{F}(t)/\bar{F}(x)dt$ for $x < x_\star$. The standard representation for these three limiting distributions, due to Mises (1936) and Jenkinson (1955), is as follows.

Definition 2.10 (Generalized Extreme Value (GEV) distribution)**.** *For* $\xi \in \mathbb{R}$, *the GEV distribution is defined for* $1 + \xi x > 0$ *by*

$$H_\xi(x) = \begin{cases} \exp\{-(1 + \xi x)^{-1/\xi}\} & \text{for} \quad \xi \neq 0 \\ \exp\{-\exp(-x)\} & \text{for} \quad \xi = 0 \end{cases} .$$

Clearly, the shape parameter ξ in the above definition corresponds to $1/\alpha$ in the previous theorem. GEV distributions are exactly the distributions which are max-stable, that is, distributions F such that for all $n \geq 2$, there exists $c_n > 0$ and $d_n \in \mathbb{R}$ so that

$$\max\{X_1, \ldots, X_n\} \stackrel{d}{=} c_n X + d_n$$

where $X_1, \ldots, X_n$ are independent and identically distributed according to F. See Theorem 3.2.2 in Embrechts et al. (1997) for example.

In the multivariate setting, consider an i.i.d. sample $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n$ from a d-dimensional distribution F with marginals $F_1, \ldots, F_d$. Define the component wise maxima $M_{jn} = \max\{X_{j1} \ldots, X_{jn}\}$ for $j \in \{1, \ldots, d\}$. Suppose that there exists sequences $a_{jn} > 0$ and $b_{jn} \in \mathbb{R}$, $j \in \{1, \ldots, d\}$, such that for $\boldsymbol{x} = (x_1, \ldots, x_d) \in \mathbb{R}^d$,

$$\left(\frac{M_{1n} - b_{1n}}{a_{1n}}, \ldots, \frac{M_{dn} - b_{dn}}{a_{dn}} \right) \rightsquigarrow G, \tag{2.7}$$

as $n \to \infty$ for some non-degenerate G where $\rightsquigarrow$ denotes convergence in distribution. If this is the case, then these margins are GEV by Theorem 2.4. Moreover, G is called a multivariate extreme-value distribution (MEV) and $F \in \mathcal{M}(G)$. Since the margins are continuous, Theorem 2.1 guarantees the existence of a unique copula for G. Analogously to the univariate setting, G must be max-stable, which is the case if and only if its margins are GEV and its copula C is extreme-value (see Theorem 7.44 in McNeil et al. (2005)).

Theorem 2.5. *If* (2.7) *holds for some G with GEV margins, then the unique copula C of G must be extreme-value. That is, for all $\boldsymbol{u} \in [0,1]^d$ and all $t \geq 0$:*

$$C(\boldsymbol{u}) = C^t(\boldsymbol{u}^{1/t}) \, .$$

There are many mathematical characterizations of MEV distributions. In this thesis, the characterization of MEVs by stable tail dependence functions is the most convenient approach. They were first introduced by Huang (1992).

Definition 2.11. *A function $\ell : \mathbb{R}_+^d \to \mathbb{R}^+$ is called a d-variate stable tail dependence function (stdf) if there exists a finite measure H on the d-dimensional unit simplex Δ_d such that for all $j \in \{1, \dots, d\}$, $\int_{\Delta_d} s_j dH(\boldsymbol{s}) = 1$, and such that for all $\boldsymbol{x} \in \mathbb{R}_+^d$,*

$$\ell(\boldsymbol{x}) = \int_{\Delta_d} \max(x_1 s_1, \dots, x_d s_d) dH(\boldsymbol{s}).$$

Stable tail dependence functions are fully characterized by Ressel (2013) as follows.

Theorem 2.6. *$\ell : \mathbb{R}_+^d \to \mathbb{R}^+$ is a d-variate stdf if and only if*

(a) *ℓ is homogeneous of degree 1, i.e., for all $k > 0$ and $x_1, \dots, x_d \in [0, \infty)$, $\ell(kx_1, \dots, kx_d)$ $= k\,\ell(x_1, \dots, x_d)$;*

(b) *$\ell(\boldsymbol{e}_1) = \cdots = \ell(\boldsymbol{e}_d) = 1$ where for $j \in \{1, \dots, d\}$, $\boldsymbol{e}_j$ denotes a vector whose components are all 0 except the jth which is equal to 1;*

(c) *ℓ is fully d-max decreasing, i.e., for any $k \in \mathbb{N}$, $x_1, \dots, x_d, h_1, \dots, h_d \in [0, \infty)$ and $J \subseteq \{1, \dots, d\}$ with $|J| = k$,*

$$\sum_{\iota_1, \dots, \iota_k \in \{0,1\}} (-1)^{\iota_1 + \cdots + \iota_k} \ell(x_1 + \iota_1 h_1 \mathbf{1}_{1 \in J}, \dots, x_d + \iota_d h_d \mathbf{1}_{d \in J}) \leq 0.$$

With the notion of stable tail dependence functions, we can now characterize extreme-value copulas.

Theorem 2.7. *A copula C is extreme-value if and only if there exists a stable tail dependence function such that $C = C_\ell$, where for all $\boldsymbol{u} \in [0,1]^d$*

$$C_\ell(\boldsymbol{u}) = \exp\{-\ell(-\log u_1, \dots, -\log u_d)\} \, .$$

Another characterization of extreme-value copulas, initially proposed in the bivariate setting only, relies on the so-called Pickands dependence function denoted A and due to Pickands (1981). Due to homogeneity (Property (a) in Theorem 2.6), an stdf ℓ is uniquely determined by its restriction A to the unit simplex via $\ell(\boldsymbol{x}) = ||\boldsymbol{x}||A(\boldsymbol{x}/||\boldsymbol{x}||)$, $\boldsymbol{x} \in \mathbb{R}^d_+$. For pair of random variables X_1, X_2 with bivariate extreme-value copula $C_\ell = C_A$, the Pickands dependence function A is in fact defined on $[0,1]$ and it is easily shown that

$$\lambda_U(X_1, X_2) = 2 - 2A(1/2) = 2 - \ell(1,1) \ , \qquad \lambda_L(X_1, X_2) = 0 \ .$$

$$\eta_U(X_1, X_2) = \begin{cases} 1/2 & \text{if} \quad A(1/2) = 1 \\ 1 & \text{otherwise} \end{cases} , \qquad \eta_L(X_1, X_2) = 1/(2A(1/2)) = 1/\ell(1,1) \ .$$

Note that if $A(1/2) = 1$, then by convexity $A(t) = 1$ for all $t \in [0,1]$ so that $C_A = C_\Pi$. Moreover, Kendall's tau has can be written in integral form viz. $\tau(X_1, X_2) = \int_0^1 \{t(1-t)/A(t)\} dA'(t)$, as shown by Ghoudi et al. (1998).

Weakening the independence assumption on the convergence to extreme value distributions is of course desirable and still an active area of research today. Leadbetter et al. (1983) established the so called $D(u_n)$ and $D'(u_n)$ conditions on temporal dependence for the univariate theory and Hsing (1989); Hüsler (1990) studied the multivariate setting using beta-mixing (stronger than the alpha-mixing to be introduced in Section 2.2).

We can now state conditions under which $F \in \mathcal{M}(G)$ and define the so-called copula domain of attraction.

Theorem 2.8. *Let $F(\boldsymbol{x}) = C(F_1(x_1), \ldots, F_d(x_d))$ for continuous marginal distribution functions $F_1, \ldots, F_d$ and some copula C. Let $G(\boldsymbol{x}) = C_0(G_1(x_1), \ldots, G_d(x_d))$be an MEV distribution with extreme-value copula C_0. Then $F \in \mathcal{M}(G)$ if and only if $F_j \in \mathcal{M}(G_j)$ for $j \in \{1, \ldots, d\}$ and for all $\boldsymbol{u} \in [0,1]^d$,*

$$\lim_{t \to \infty} C^t(u_1^{1/t}, \ldots, u_d^{1/t}) = C_0(u_1, \ldots, u_d) \ .$$

Moreover, we say that C is in the copula domain of attraction of C_0, written $C \in CDA(C_0)$.

2.1.4 The Archimax family

The class of so-called Archimax copulas was proposed by Capéraà et al. (2000) in the bivariate case and extended to higher dimensions by Mesiar and Jágr (2013) and Charpentier et al. (2014). The latter are, at any $\boldsymbol{u} \in [0,1]^d$, of the form

$$C_{\psi,\ell}(\boldsymbol{u}) = \psi[\ell\{\phi(u_1), \ldots, \phi(u_d)\}], \tag{2.8}$$

where ℓ is an arbitrary d-variate stdf and $\psi : [0,\infty) \to [0,1]$ is an Archimedean generator with inverse ϕ, as in Definition 2.5. One can think of the function ψ as distorting the extreme-value dependence structure. Indeed, if $\psi(x) = e^{-x}$, then $C_{\psi,\ell} = C_\ell$ is an extreme-value copula.

The density of an Archimax copula $c_{\psi,\ell}$ can be obtained with the application of Faà di Bruno's formula, as shown for example by Hofert et al. (2018). It can be written, for all $\boldsymbol{u} \in (0,1)^d$, as

$$c_{\psi,\ell}(\boldsymbol{u}) = \left\{\prod_{j=1}^{d} \phi'(u_j)\right\} \sum_{k=1}^{d} \psi^{(k)}[\ell\{\phi(\boldsymbol{u})\}] \sum_{\pi\in\Pi:|\pi|=k} \prod_{B\in\pi} (D_B\ell)\{\phi(\boldsymbol{u})\} ,$$

where $D_B\ell$ denotes the partial derivatives of ℓ with respect to the variables in the index set B and Π denotes the set of all partitions of $\{1,\ldots,d\}$. We begin with a definition of several key concepts including Archimax copulas.

Definition 2.12. *A d-dimensional copula C is called Archimax if it permits the representation* (2.8) *for some d-variate stdf ℓ and an Archimedean generator ψ with inverse ϕ as defined in Definition 2.5.*

As the name suggests, the class of Archimax copulas includes both Archimedean and extreme-value copulas. When ℓ is the stdf pertaining to independence, i.e., $\ell(\boldsymbol{x}) = x_1 + \cdots + x_d$ for all $\boldsymbol{x} \in \mathbb{R}_+^d$, $C_{\psi,\ell}$ in (2.8) becomes the Archimedean copula C_ψ with generator ψ. When $\psi(x) = e^{-x}$ for any $x \geq 0$, $C_{\psi,\ell}$ reduces to the extreme-value copula C_ℓ with stdf ℓ. An interesting special case arises when $\ell = \ell_M$ with $\ell_M(\boldsymbol{x}) = \max(x_1,\ldots,x_d)$ for all $\boldsymbol{x} \in \mathbb{R}_+^d$. Because ϕ is strictly decreasing on $(0,1]$, one has that for all $\boldsymbol{u} \in [0,1]^d$, $C_{\psi,\ell_M}(\boldsymbol{u}) = \min(u_1,\ldots,u_d)$. In other words, C_{ψ,ℓ_M} is the Fréchet–Hoeffding upper bound whatever the generator ψ; this copula characterizes the dependence between comonotonic variables.

The right-hand side in (2.8) is not a bona fide copula for all choices of Archimedean generators and d-variate stdfs and d-variate stdf. As proved by Charpentier et al. (2014), a sufficient condition is that ψ is d-monotone. When $\ell(\boldsymbol{x}) = x_1 + \cdots + x_d$, i.e., when $C_{\psi,\ell}$ is Archimedean, the d-monotonicity of ψ is also necessary as discussed by Malov (2001); Morillas (2005); McNeil and Nešlehová (2009). However, this condition is not necessary in general; Example 3.7 of Charpentier et al. (2014) shows that for some stdfs, it suffices that ψ is k-monotone for some $k < d$. In fact, ψ can be an arbitrary Archimedean generator when $\ell = \ell_M$.

An Archimax copula can also be defined $C_{\psi,A}$, i.e. in terms of a Pickands dependence function instead of an stdf, and expressed, for any $\boldsymbol{u} \in [0,1]^d$, as

$$C_{\psi,A}(\boldsymbol{u}) = \psi\left[\|\phi(\boldsymbol{u})\| A\left\{\phi(\boldsymbol{u})/\|\phi(\boldsymbol{u})\|\right\}\right]. \tag{2.9}$$

Archimax copulas admit a stochastic representation similar to that of Archimedean copulas. Let R be a nonnegative random variable, with distribution F_R, independent of $\boldsymbol{S}_d$, a random vector with survival function defined, for $\boldsymbol{s} \in \mathbb{R}_+^d$, by

$$\Pr(S_1 > s_1,\ldots,S_d > s_d) = \left[\max\{0, 1-\ell(\boldsymbol{s})\}\right]^{d-1} , \tag{2.10}$$

where ℓ is a stable tail dependence function. As was the case for Archimedean copulas, we are interested in the survival copula of vectors of the form

$$\boldsymbol{X} = R\boldsymbol{S}_d = R \times (S_1, \ldots, S_d) , \tag{2.11}$$

but here $\boldsymbol{S}$ belongs to a larger class of distributions. This stochastic representation is formally shown in Theorem 3.3 from Charpentier et al. (2014). In this representation, R can again be interpreted as a distortion variable; when its law is Erlang with parameter d, $C_{\psi,\ell} = C_\ell$.

Theorem 2.9. *(i) If $(X_1, \ldots, X_d)$ is a random vector of the form (2.11), then its survival copula is the Archimax copula $C_{\psi,\ell}$, where ψ is the Williamson d-transform of F_R.*

(ii) Let ℓ be a d-variate stable tail dependence function anf ψ be a generator of a d-dimensional Archimedean copula. Then $C_{\psi,\ell}$ is the survival copula of a random vector of the form (2.11), where the distribution function F_R is the inverse Williamson d-transform of ψ.

The Archimax copulas have given extreme-value attractors; Propositions 6.1 and 6.4 from Charpentier et al. (2014), regarding the maximum and minimum domains of attraction respectively, are reproduced in the following.

Proposition 2.1. *Suppose that ψ is a generator of a d-variate Archimedean copula with $1 - \psi(1/\cdot) \in \mathcal{R}_{-\alpha}$ for some $\alpha \in (0, 1]$. Then the Archimax copula $C_{\psi,\ell}$ belongs to the copula domain of attraction of the extreme-value copula C_{ℓ_α} where for all $\boldsymbol{x} \in \mathbb{R}^d_+$,*

$$\ell_\alpha(\boldsymbol{x}) = \ell^\alpha(\boldsymbol{x}^{1/\alpha}) .$$

Equivalently, $\lim_{n\to\infty} C^n_{\psi,\ell}(\boldsymbol{u}^{1/n}) = C_{\ell_\alpha}(\boldsymbol{u})$.

Remark 2.3. *It is clear from this result that Archimedean copulas belong to the copula domain of attraction of the Gumbel (or logistic) family. Indeed as noted earlier, if $\ell(\boldsymbol{x}) = x_1 + \cdots + x_d$, then $C_{\psi,\ell} = C_\psi$ and provided $1 - \psi(1/\cdot) \in \mathcal{R}_{-\alpha}$, $\lim_{n\to\infty} C^n_\psi(\boldsymbol{u}^{1/n}) = C_{\ell_\alpha}(\boldsymbol{u})$ where $\ell_\alpha(\boldsymbol{x}) = (x_1^{1/\alpha} + \cdots + x_d^{1/\alpha})^\alpha$, the logistic stable tail dependence function. This was initially proved by Genest and Rivest (1989).*

Suppose $\boldsymbol{X}$ has copula C. To find the minimum domain of attraction of C, the variable of interest is the component-wise minimum, i.e. $\boldsymbol{W}_n = (W_{1n}, \ldots, W_{dn})$ where $W_{jn} = \min\{X_{j1}, \ldots, X_{jn}\}$, $j = 1, \ldots, d$. Using (2.1) and elementary algebra, one has that for $\boldsymbol{u} \in (0, 1)^d$, the copula of $\boldsymbol{W}_n$ is as follows:

$$C_{\boldsymbol{W}_n}(\boldsymbol{u}) = \sum_{\iota_1,\ldots,\iota_d \in \{0,1\}} (-1)^{\iota_1+\ldots+\iota_d} \bar{C}^n\{(1 - \iota_1 v_1)^{1/n}, \ldots, (1 - \iota_d v_d)^{1/n}\} ,$$

where $\bar{C}$ is the survival copula of C. The following proposition by Charpentier et al. (2014) determines the limit of $C_{\boldsymbol{W}_n}$ as $n \to \infty$ when C is Archimax.

Proposition 2.2. *Suppose that ψ is the generator of a d-variate Archimedean copula with $\phi(1/\cdot) \in \mathcal{R}_\alpha$ for some $\alpha \in (0,\infty)$. Then the survival copula of $C_{\psi,\ell}$, denoted $\bar{C}_{\psi,\ell}$, is in the copula domain of attraction of $D^\star$ defined, for all $\boldsymbol{u} \in (0,1)^d$ by*

$$D^\star(\boldsymbol{u}) = \sum_{\iota_1,\ldots,\iota_d \in \{0,1\}} (-1)^{\iota_1+\ldots+\iota_d} K(\iota_1 u_1, \ldots, \iota_d u_d) ,$$

where for arbitrary $v_1, \ldots, v_d \in [0,1]$,

$$K(v_1,\ldots,v_d) = \exp\left\{ - \sum_{\iota_1,\ldots,\iota_d \in \{0,1\}} (-1)^{\iota_1+\ldots+\iota_d} \ln C_{\psi^\star,\ell}(1-\iota_1 v_1, \ldots, 1-\iota_d v_d) \right\} ,$$

with $\psi^\star(t) = \exp(-t^{-1/\alpha})$ for all $t > 0$. That is, $\bar{C}_{\psi,\ell} \in CDA(D^\star)$. This is also equivalent to saying that the Archimax copula $C_{\psi,\ell}$ belongs to the minimum domain of attraction of an extreme-value distribution whose unique underlying copula is $D^\star$.

It is clear from the two previous propositions that the stable tail dependence function ℓ of an Archimax copula is the main driver of its extreme behavior. However, the regular variation of the generator also plays a role. This regular variation translates to tail behavior of the radial variable R in the stochastic representation in (2.11). Indeed, Theorem 2 from Larsson and Nešlehová (2011) shows that $1-\psi(1/\cdot) \in \mathcal{R}_{-\alpha}$ if and only if $1/R \in \mathcal{M}(\Phi_\alpha)$ for $\alpha \in (0,1)$. Moreover, $1-\psi(1/\cdot) \in \mathcal{R}_{-1}$ if $1/R$ is in the maximum domain of attraction of the Weibull distribution, Gumbel distribution or Fréchet distribution with $\alpha \geq 1$. For the minimum attractor, the condition that $\phi(1/\cdot) \in \mathcal{R}_\alpha$ occurs if and only if $R \in \mathcal{M}(\Phi_{1/\alpha})$. Bücher et al. (2019) have linked these indices of regular variation with the speed of convergence of Archimax copulas to their extreme-value attractor.

2.2 The empirical copula process

This section defines the convergence concepts used later in Chapters 4 and 5. However, only the essential and necessary elements to understand the derived asymptotic results are summarized here.

2.2.1 Weak convergence

This section is based on the text by van der Vaart and Wellner (1996), in which one will find a much deeper view into this area. For a metric space (D,d), let $(P_n)_{n=1}^\infty$ and P be Borel probability measures defined on $(D,\mathcal{D})$, where $\mathcal{D}$ is a Borel σ-algebra on D. Let $C_b(D)$ denote the space of bounded, continuous, real functions on D. The sequence P_n is said to converge weakly to P if and only if for all $f \in C_b(D)$,

$$\int_D f dP_n \to \int_D f dP$$

as $n \to \infty$. This is denoted by $P_n \rightsquigarrow P$. The equivalent definition of weak convergence for D-valued random variables $(X_n)_{n=1}^{\infty}$ and X is that $X_n \rightsquigarrow X$ if and only if for all $f \in C_b(D)$,

$$\mathrm{E}\, f(X_n) \to \mathrm{E}\, f(X) \tag{2.12}$$

as $n \to \infty$. Classically, the theory requires that for each n, P_n is defined on the Borel σ-field $\mathcal{D}$, which is equivalent to saying that X_n is Borel measurable. If D is separable this condition usually holds but if it is non-separable it can sometimes fail. For example, it holds for $C[0,1]$ (the space of continuous functions on $[0,1]$) with the supremum norm, but it fails on $D[0,1]$ (the Skohorod space of càdlàg functions on $[0,1]$) with the supremum norm.

Pursuing the latter example, let $U_1, \ldots, U_n$ be independent random variables uniformly distributed on $[0,1]$. Now let the empirical distribution function F_n be defined for $u \in [0,1]$ as

$$F_n(u) = \frac{1}{n} \sum_{i=1}^{n} \mathbf{1}(U_i \leq u)$$

and the uniform empirical process, for $u \in [0,1]$, as

$$X_n(u) = \sqrt{n}(F_n(u) - u) \ .$$

Both F_n and X_n are maps from $[0,1]^n$ to $D[0,1]$. However, neither is Borel measurable if $D[0,1]$ is equipped with the supremum norm. Out of all possible solutions to alleviate this, the monograph from van der Vaart and Wellner (1996) focuses on the notion of outer expectation and probability as proposed by Hoffman-Jorgensen (1994).

Definition 2.13. *Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space and $T : \Omega \to \bar{\mathbb{R}}$ an arbitrary map.*

- *The outer expectation of T with respect to $\mathbb{P}$ is defined as*

$$\mathbb{E}^{\star} T = \inf\{\mathbb{E}U : U \geq T, U : \Omega \to \bar{\mathbb{R}} \textit{ measurable and } \mathbb{E}U \textit{ exists}\} \ ,$$

  where $\mathbb{E}U$ is understood to exist if $\mathbb{E}|U|$ exists.

- *The outer probability of any $B \subset \Omega$ is defined as*

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

Inner expectation and probability are then easy to define as well.

Definition 2.14. *Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space and $T : \Omega \to \bar{\mathbb{R}}$ an arbitrary map.*

- *The inner expectation of T with respect to $\mathbb{P}$ is defined as*

$$\mathbb{E}_{\star} T = -\mathbb{E}^{\star}\{-T\} \ ,$$

- *The inner probability of any $B \subset \Omega$ is defined as*

$$\mathbb{P}_\star(B) = 1 - \mathbb{P}^\star(B^c)\ .$$

We now have the ingredients to define weak convergence for possibly non Borel-measurable maps.

Definition 2.15. *Let $X_n : \Omega_n \to D$, $n \in \mathbb{N}$ and $X : \Omega \to D$ be arbitrary maps from the probability spaces $(\Omega_n, \mathcal{A}_n, \mathbb{P}_n)$ and $(\Omega, \mathcal{A}, \mathbb{P})$ respectively. Let X be Borel-measurable. The sequence X_n converges weakly to X, that is $X_n \rightsquigarrow X$, if for any $f \in C_b(D)$,*

$$\mathbb{E}^\star\{f(X_n)\} \to \mathbb{E}\{f(X)\}$$

as $n \to \infty$.

Many tools available to the classical concept of weak convergence such as the continuous mapping theorem are available for this concept as well (as shown by van der Vaart and Wellner (1996)). Inner expectation allows us to define asymptotic measurability and tightness.

Definition 2.16. *Let $(\Omega_n, \mathcal{A}_n, \mathbb{P}_n)$ be arbitrary probability spaces and $X_n : \Omega_n \to D$ be arbitrary mappings, $n \in \mathbb{N}$. The sequence $(X_n)_{n=1}^\infty$ is asymptotically measurable if and only if $\mathbb{E}^\star\{f(X_n)\} - \mathbb{E}_\star\{f(X_n)\} \to 0$ as $n \to \infty$ for all $f \in C_b(D)$.*

Definition 2.17. *Let $(\Omega_n, \mathcal{A}_n, \mathbb{P}_n)$ be arbitrary probability spaces and $X_n : \Omega_n \to D$ be arbitrary mappings, $n \in \mathbb{N}$. The sequence $(X_n)_{n=1}^\infty$ is asymptotically tight if and only if for any $\epsilon > 0$, there exists a compact set $K \subset D$ such that $\liminf_{n\to\infty} \mathbb{P}_\star(X_n \in O) \geq 1 - \epsilon$ for any open set $O \supset K$.*

The following Lemma 1.3.8 from van der Vaart and Wellner (1996) shows the connection between weak convergence and asymptotic measurability and tightness.

Lemma 2.1.
- *If $X_n \rightsquigarrow X$ as $n \to \infty$ then $(X_n)_{n=1}^\infty$ is asymptotically measurable.*
- *If $X_n \rightsquigarrow X$ as $n \to \infty$, then $(X_n)_{n=1}^\infty$ is asymptotically tight if and only if X is tight.*

To get to an intuitive notion of weak convergence using asymptotic tightness, we restrict ourselves to spaces of uniformly bounded functions. The following result is proved in Theorem 1.5.4 from van der Vaart and Wellner (1996). Recall that for an arbitrary domain S, $\ell^\infty(S)$ denotes the space of functions $f : S \to \mathbb{R}$ such that $||f||_\infty = \sup_{s\in S} |f(s)| < \infty$ equipped with the supremum norm.

Theorem 2.10. *Suppose that $X_n : \Omega_n \to \ell^\infty(S)$, $n \in \mathbb{N}$, are arbitrary maps. Then $X_n \rightsquigarrow X$ in $\ell^\infty(S)$ if and only if*

- $(X_n(s_1), \ldots, X_n(s_k))$ *converges weakly to* $(X(s_1), \ldots, X(s_k))$ *in* $\mathbb{R}^k$ *for any finite subset* $s_1, \ldots, s_k$ *of* S.
- $(X_n)_{n=1}^{\infty}$ *is asymptotically tight.*

Equivalently, the second condition in the above theorem can be replaced by asymptotic uniform equicontinuity as defined below.

Definition 2.18. *Let* $X_n : \Omega_n \mapsto \ell^\infty(S)$, $n \in \mathbb{N}$, *be arbitrary maps. The collection* $(X_n)_{n=1}^{\infty}$ *is asymptotically uniformly equicontinuous in probability with respect to a semi-metric* ρ *if and only if, for every* $\epsilon, \eta > 0$ *there exists a* $\delta > 0$ *such that*

$$\limsup_{n\to\infty} \mathbb{P}^\star\Big(\sup_{\rho(s,t)<\delta} |X_n(s) - X_n(t)| > \epsilon\Big) < \eta \ .$$

2.2.2 The empirical process

The empirical measure of a sample of random variables, as introduced below, is simply a linear combination of Dirac measures at the observations, each with weight $1/n$.

Definition 2.19. *Let* $X_1, \ldots, X_n$ *be a random sample in the measurable space* $(\mathcal{X}, \mathcal{A})$. *The empirical measure of* $X_1, \ldots, X_n$ *is defined for any* $A \in \mathcal{A}$ *as*

$$\mathbb{P}_n(A) = \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}(X_i \in A) \ .$$

For any signed measure Q and a measurable function $f : \mathcal{X} \to \mathbb{R}$, let $Qf = \int f dQ$. For a collection $\mathcal{F}$ of such measurable functions, an empirical measure $\mathbb{P}_n$ induces a map from $\mathcal{F}$ to $\mathbb{R}$ by

$$f \mapsto \mathbb{P}_n f \ .$$

Definition 2.20. *Let* $X_1, \ldots, X_n$ *be a random sample in the measurable space* $(\mathcal{X}, \mathcal{A})$ *with common distribution* P *and* $\mathcal{F}$ *be a collection of measurable functions* $f : \mathcal{X} \to \mathbb{R}$. *The empirical process of* $X_1, \ldots, X_n$ *indexed by* $\mathcal{F}$ *is defined as the following rescaled and centered map*

$$f \mapsto \mathbb{G}_n f = \sqrt{n}(\mathbb{P}_n - P)f = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}(f(X_i) - Pf) \ .$$

The classical empirical process is obtained by simply restricting the sample space $\mathcal{X}$ to be $[0,1]$, $\mathbb{R}$, $[0,1]^d$ or $\mathbb{R}^d$ and $\mathcal{F}$ to be the collection of indicator functions of left half-lines (or lower-left orthants of $\mathbb{R}^d$).

Glivenko-Cantelli and *Donsker* classes can now be defined. The uniform version of the law of large numbers becomes

$$||\mathbb{P}_n - P||_{\mathcal{F}} \to 0$$

as $n \to \infty$, where $||Q||_{\mathcal{F}} = \sup\{|Qf| : f \in \mathcal{F}\}$ and the convergence is either in outer probability or outer almost surely. A class $\mathcal{F}$ for which this is true is called a *P-Glivenko-Cantelli* class. To consider a uniform version of the central limit theorem, one needs to assume that for all x,

$$\sup_{f \in \mathcal{F}} |f(x) - Pf| < \infty .$$

This implies that $\mathbb{G}_n \in \ell^\infty(\mathcal{F})$. Under assumptions on $\mathcal{F}$, one can show that

$$\mathbb{G}_n = \sqrt{n}(\mathbb{P}_n - P) \rightsquigarrow \mathbb{G}$$

in $\ell^\infty(\mathcal{F})$, for some Borel-measurable and tight limit $\mathbb{G} \in \ell^\infty(\mathcal{F})$. A class of functions $\mathcal{F}$ for which this holds is called *P-Donsker*. These conditions for $\mathcal{F}$ to be P-Donsker, namely bounded uniform (or bracketing, alternatively) entropy, are discussed in Chapter 2.5 in van der Vaart and Wellner (1996). Clearly, a *Donsker* class is also a *Glivenko-Cantelli* class but the converse is not always true. Naturally, one would want to know more about the limiting process $\mathbb{G}$. Firstly, the marginals $\mathbb{G}_n f$ converge if and only if f are square-integrable. If this holds, the multivariate central limit theorem implies that for any finite set $f_1, \ldots, f_k$,

$$(\mathbb{G}_n f_1, \ldots, \mathbb{G}_n f_k) \rightsquigarrow N(0, \Sigma) ,$$

where $N(0, \Sigma)$ is a k-dimensional standard normal distribution whose variance-covariance matrix Σ has (i, j)-th entry $P(f_i - Pf_i)(f_j - Pf_j)$. It follows that $\{\mathbb{G}f : f \in \mathcal{F}\}$ is a zero-mean Gaussian process with covariance

$$\mathrm{E}\, \mathbb{G}f_1 \mathbb{G}f_2 = P(f_1 - Pf_1)(f_2 - Pf_2) = Pf_1 f_2 - Pf_1 Pf_2 .$$

Due to its tightness, Lemma 1.5.3 from van der Vaart and Wellner (1996) ensures that the distribution of $\mathbb{G}$ in $\ell^\infty(\mathcal{F})$ is completely determined by the above covariance function. $\mathbb{G}$ is called a P-Brownian bridge (or sheet when the dimension of $\mathcal{X}$ is larger than 1).

2.2.3 Weak convergence of the empirical copula process

Now that notions regarding weak convergence and empirical processes are defined, emphasis is made on the specific case of the empirical copula process. Let $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n$ be an independent and identically distributed (i.i.d.) sample from a d-dimensional distribution F with continuous marginal distribution functions $F_1, \ldots, F_d$ and unknown copula C. A natural and non-parametric estimate of each marginal distribution is the so-called empirical distribution function given for $j \in \{1, \ldots, d\}$ and $x \in \mathbb{R}$ by $F_{nj}(x) = n^{-1} \sum_{i=1}^n \mathbf{1}(X_{ij} \le x)$. This can be naturally extended to the multivariate setting by letting, for $x \in \mathbb{R}^d$, $F_n(x) = n^{-1} \sum_{i=1}^n \mathbf{1}(\boldsymbol{X}_i \le \boldsymbol{x})$. To construct the empirical copula, first define the normalized ranks as follows, for $i \in \{1, \ldots, n\}$ and $j \in \{1, \ldots, d\}$,

$$\hat{U}_{ij} = n\, F_{nj}(X_{ij})/(n+1) . \tag{2.13}$$

The rank-based empirical copula can now be defined for $\boldsymbol{u} \in [0,1]^d$ by

$$\hat{C}_n(\boldsymbol{u}) = \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}(\hat{\boldsymbol{U}}_i \leq \boldsymbol{u}) = \frac{1}{n}\sum_{i=1}^{n}\prod_{j=1}^{d} \mathbf{1}(\hat{U}_{ij} \leq u_j) \ . \tag{2.14}$$

The above is simply the empirical distribution of the renormalized ranks of the observed data (see results from Rüschendorf (1976) for example). This empirical copula is slightly different than as first introduced by Deheuvels (1979), for $\boldsymbol{u} \in [0,1]^d$,

$$C_n(\boldsymbol{u}) = F_n(F_{n1}^{-1}(u_1), \ldots, F_{nd}^{-1}(u_d)) \ . \tag{2.15}$$

Note that neither expressions (2.15) nor (2.14) is a copula *stricto sensu*. Also note that (2.15) and (2.14) are asymptotically equivalent. It is even shown in Lemma 4.6 by Berghaus et al. (2017) that $\sup_{\boldsymbol{u}\in[0,1]^d} |C_n(\boldsymbol{u}) - \hat{C}_n(\boldsymbol{u})| = o_p(n^{-1/2})$ for weakly dependent samples. The empirical copula process is then simply defined for all $\boldsymbol{u} \in [0,1]^d$ as

$$\hat{\mathbb{C}}_n(\boldsymbol{u}) = \sqrt{n}\{\hat{C}_n(\boldsymbol{u}) - C(\boldsymbol{u})\} \ . \tag{2.16}$$

Of course, replacing $\hat{C}_n$ by C_n in the above does not change the limit of the process. This asymptotic behavior has been the subject of many papers over the years. Overall, smoothness conditions on C have been weakened and convergence results are also now available under certain serial dependence conditions on $\boldsymbol{X}_1, \boldsymbol{X}_2, \ldots$. First, the limiting distribution under independence of the margins was established by Deheuvels (1981a,b). Weak convergence in the Skorohod space $D([0,1]^d)$ was established by Rüschendorf (1976) and Gaenssler and Stute (1987), with less restrictive assumptions in the latter. One can also refer to Example 3.9.29 in van der Vaart and Wellner (1996) for another convergence result in the Skorohod space restricted to a closed set in the interior of $[0,1]^2$. Weak convergence in $\ell^\infty([0,1]^d)$ was established by Fermanian et al. (2004) with conditions on the first order derivatives of C. Convergence rates were proposed by Stute (1984) and studied by Tsukahara (2000) under assumptions on second order derivatives. These assumptions were then weakened by Segers (2012). In the work of Bücher and Volgushev (2013), convergence is established for weak serial dependence of the sample, a much more realistic condition than serial independence. To allow for broader applications, Berghaus et al. (2017) proved weak convergence of the empirical process with respected to stronger weighted metrics. Convergence of the empirical copula process was also studied in the case where the underlying distributions lack a certain degree of smoothness. Genest et al. (2017) study the asymptotic behavior of the empirical copula process under broad conditions that include, for example, discrete margins. Weak convergence with respect to a metric related to epi- and hypo-convergence is established by Bücher et al. (2014). The asymptotic behavior of the estimators proposed in this thesis is established thanks to the work of Berghaus et al. (2017). Firstly, the following notion of asymptotic serial independence allows to relax serial independence.

Definition 2.21. *For $-\infty \leq a < b \leq \infty$, let $\mathcal{F}_a^b$ be the σ-field generated by $(\boldsymbol{X}_i)_{i\in\mathbb{Z}}$ with $i \in \{a, a+1, \ldots, b\}$. For $k \geq 1$, define $\alpha^{[\boldsymbol{X}]}(k) = \sup\{|\Pr(A \cap B) - \Pr(A)\Pr(B)| : A \in \mathcal{F}_{-\infty}^i, B \in \mathcal{F}_{i+k}^{\infty}, i \in \mathbb{Z}\}$ as the alpha-mixing coefficient of $(\boldsymbol{X}_i)_{i\in\mathbb{Z}}$. The series is called alpha-mixing (or strongly mixing) if $\alpha^{[\boldsymbol{X}]}(k) \to 0$ as $k \to \infty$.*

Next, smoothness assumptions on the true copula C are needed.

Condition 2.1. *For $l \in \{1, \ldots, d\}$, let $V_{d,l} = \{\boldsymbol{u} \in [0,1]^d : u_l \in (0,1)\}$. For each $j \in \{1, \ldots, d\}$, the partial derivative $\dot{C}_j$ given for all $\boldsymbol{u} \in [0,1]^d$ by $\dot{C}_j(\boldsymbol{u}) = \partial C(\boldsymbol{u})/\partial u_j$ exists and is continuous on the set $V_{d,j}$.*

For a d-variate copula C, let α be a C-Brownian bridge, i.e., a tight, centered Gaussian process with covariance function given, for all $\boldsymbol{u}, \boldsymbol{v} \in [0,1]^d$ by

$$\operatorname{cov}\{\alpha(\boldsymbol{u}), \alpha(\boldsymbol{v})\} = \sum_{i\in\mathbb{Z}} \operatorname{cov}\{\mathbf{1}(\boldsymbol{U}_0 \leq \boldsymbol{u}), \mathbf{1}(\boldsymbol{U}_i \leq \boldsymbol{v})\}\,, \tag{2.17}$$

where $\boldsymbol{U}_i = (F_1(X_{i1}), \ldots, F_d(X_{id}))$. Note that in the case of serial independence, this covariance function simplifies to $\operatorname{cov}\{\alpha(\boldsymbol{u}), \alpha(\boldsymbol{v})\} = C(\boldsymbol{u} \wedge \boldsymbol{v}) - C(\boldsymbol{u})C(\boldsymbol{v})$. Finally, let $\mathbb{C}$ be the process defined, for any $\boldsymbol{u} \in [0,1]^d$, by

$$\mathbb{C}(\boldsymbol{u}) = \alpha(\boldsymbol{u}) - \sum_{j=1}^{d} \dot{C}_j(\boldsymbol{u})\alpha(\boldsymbol{u}^{(j)}) \tag{2.18}$$

with $\boldsymbol{u}^{(j)} = (1, \ldots, 1, u_j, 1, \ldots, 1)$. For any $j \in \{1, \ldots, d\}$ and $\boldsymbol{u} \in [0,1]^d$, if the derivative $\partial C(\boldsymbol{u})/\partial u_j$ does not exist, set $\dot{C}_j(\boldsymbol{u}) = \limsup_{h\to 0}\{C(\boldsymbol{u} + h\boldsymbol{e}_j) - C(\boldsymbol{u})\}$. The following result was proved by Bücher and Ruppert (2013) (Theorem 1).

Theorem 2.11. *If Condition 2.1 holds and if $(\boldsymbol{X}_i)_{i\in\mathbb{Z}}$ are alpha-mixing with $\alpha^{[\boldsymbol{X}]}(k) = O(k^{-a})$ with $a > 1$, then $\hat{\mathbb{C}}_n \rightsquigarrow \mathbb{C}$ in $(\ell^\infty([0,1]^d), ||\cdot||_\infty)$.*

Define the (unobservable) empirical process based on $\boldsymbol{U}_i = (F_1(X_{i1}), \ldots, F_d(X_{id}))$, $i \in \{1, \ldots, n\}$, for any $\boldsymbol{u} \in [0,1]^d$, by

$$\alpha_n(\boldsymbol{u}) = \sqrt{n}\{G_n(\boldsymbol{u}) - C(\boldsymbol{u})\}\,, \tag{2.19}$$

where $G_n(\boldsymbol{u}) = n^{-1}\sum_{i=1}^n \mathbf{1}(\boldsymbol{U}_i \leq \boldsymbol{u})$. In the above Theorem, the copula could be estimated by G_n if the margins were known. The limit in that case is simply α without the extra terms involving the first order partial derivatives. In fact, no assumptions on C and its derivatives would be needed. As explained by Segers (2012), these extra terms encode the impact of not knowing the quantiles F_j^{-1} and replacing them with their empirical counterparts. It is not surprising to see how these 'penalty' terms depend on the sensitivity of the copula C to change in the marginals via $\dot{C}_j$. What is surprising however is that in some cases, ignoring known information about the marginal distributions can lead to a better estimation of the copula, as explored in the bivariate case by Genest and Segers (2010) and in the multivariate case in the upcoming paper from Genest et al. (2019).

The convergence results needed in this thesis required slightly more powerful tools, which is where the following condition comes into play.

Condition 2.2. *For every $i, j \in \{1, \ldots, d\}$, the second-order partial derivative $\ddot{C}_{ij}$ given for all $\boldsymbol{u} \in [0,1]^d$ by $\ddot{C}_{ij}(\boldsymbol{u}) = \partial^2 C(\boldsymbol{u})/\partial u_i \partial u_j$ exists and is continuous on the set $V_{d,j} \cap V_{d,i}$, and there exists a constant $K > 0$ such that for all $\boldsymbol{u} \in V_{d,j} \cap V_{d,i}$,*

$$|\ddot{C}_{ij}(\boldsymbol{u})| \leq K \min\left[\frac{1}{u_i(1-u_i)}, \frac{1}{u_j(1-u_j)}\right].$$

This smoothness condition was first proposed by Segers (2012), in which the almost sure convergence rate elicited by Stute (1984) is recovered. This condition, along with Condition 2.1 and alpha-mixing, is used to establish the weak convergence of the empirical copula process with respect to weighted metrics by Berghaus et al. (2017). The proof of the weak convergence of the estimators derived in this thesis hinge on their result, reproduced in the following. As alluded to, a weight function is used. For $\omega > 0$, it is defined for $\boldsymbol{u} \in [0,1]^d$ by

$$g_\omega(\boldsymbol{u}) = \min\left\{\bigwedge_{j=1}^{d} u_j, \bigwedge_{j=1}^{d}(1 - \min_{j' \neq j} u_{j'})\right\}^\omega . \tag{2.20}$$

A slight variation preventing the weight function from vanishing is also needed. For $\boldsymbol{u} \in [0,1]^d$, let

$$\tilde{g}_\omega(\boldsymbol{u}) = g_\omega(\boldsymbol{u}) + \mathbf{1}(g_\omega(\boldsymbol{u}) = 0) . \tag{2.21}$$

Theorem 2.12. *Suppose that $(\boldsymbol{X}_i)_{i \in \mathbb{Z}}$ are an alpha-mixing series with $\alpha^{[\boldsymbol{X}]}(k) = O(k^{-a})$ with $a > 1$. Suppose that the marginal distributions $F_1, \ldots, F_d$ are continuous and the underlying copula C satisfies Conditions 2.1 and 2.2. Then, for any $c \in (0,1)$ and $\omega \in (0, 1/2)$,*

$$\sup_{\boldsymbol{u} \in [c/n, 1-c/n]^d} \left| \frac{\hat{\mathbb{C}}_n(\boldsymbol{u})}{g_\omega(\boldsymbol{u})} - \frac{\bar{\mathbb{C}}_n(\boldsymbol{u})}{g_\omega(\boldsymbol{u})} \right| = o_p(1) ,$$

where for $\boldsymbol{u} \in [0,1]^d$,

$$\bar{\mathbb{C}}_n(\boldsymbol{u}) = \alpha_n(\boldsymbol{u}) - \sum_{j=1}^{d} \dot{C}_j(\boldsymbol{u}) \alpha_n(\boldsymbol{u}^{(j)}) .$$

Moreover, $\bar{\mathbb{C}}_n/\tilde{g}_\omega \rightsquigarrow \mathbb{C}/\tilde{g}_\omega$ in $(\ell^\infty([0,1]^d), ||\cdot||_\infty)$.

The restriction of the supremum in the first result to $[c/n, 1-c/n]^d$ is due to the fact that $\hat{\mathbb{C}}_n/g_\omega$ would otherwise be unbounded if the set were to be extended towards the borders of the unit hypercube.

Chapter 3

Identifiability and smoothness of the Archimax family

This chapter establishes properties of Archimax copulas that are needed for the modeling of real datasets. As shown in Section 2.1.4, this family is characterized by two functional parameters: the Archimedean generator and the stable tail dependence function. It is necessary for any inference procedure developed for Archimax copulas to be able to distinguish these two functions, and as such Section 3.1 explores the conditions under which the latter are identifiable. The inference tools developed in this thesis are justified by both simulation studies and theoretical convergence results. For the latter results, regularity assumptions are often needed in order to use powerful theorems on the asymptotic behavior of empirical copula processes. These assumptions translate to conditions on the two functional parameters of the Archimax copulas. Such conditions are stated and verified in Section 3.2.

3.1 Identifiability

In this section, we establish conditions under which ℓ and θ are identifiable when $\psi \in \Psi = \{\psi_\theta, \theta \in \mathcal{O}\}$. To accomplish this, we first consider two arbitrary d-variate Archimax copulas $C_1 = C_{\psi_1,\ell_1}$ and $C_2 = C_{\psi_2,\ell_2}$ whose generators ψ_1, ψ_2 are not necessarily from a parametric class. The lemmas below investigate the question whether $C_1 = C_2$ implies that the generators and stdfs are equal.

Lemma 3.1. *Suppose that $C_1 = C_2$ and $\psi_1 = \psi_2 = \psi$. Then $\ell_1 = \ell_2$.*

Proof. For all $\boldsymbol{u} \in [0,1]^d$, $\ell_1\{\phi(\boldsymbol{u})\} = \ell_2\{\phi(\boldsymbol{u})\}$, and since ϕ is one-to-one, $\ell_1(\boldsymbol{x}) = \ell_2(\boldsymbol{x})$ for all $\boldsymbol{x} \in \mathbb{R}_+^d$. □

Lemma 3.2. *Suppose that $C_1 = C_2$ and $\ell_1 = \ell_2 = \ell$ is a d-variate stdf such that $\ell \neq \ell_M$, where for each $\boldsymbol{x} \in \mathbb{R}_+^d$, $\ell_M(\boldsymbol{x}) = \max(x_1, \ldots, x_d)$. Suppose also that ψ_1 and ψ_2 are 2-monotone Archimedean generators. Then there exists a constant $c > 0$ such that, for all $x \geq 0$, $\psi_1(x) = \psi_2(cx)$.*

Proof. If $\ell(\boldsymbol{x}) = \ell_M(\boldsymbol{x}) = \max\{x_1, \dots, x_d\}$ for all $\boldsymbol{x} \in \mathbb{R}_+^d$, then regardless of ψ_1 and ψ_2, we have that $C_1 = C_2 = C_M$, the copula corresponding to the Fréchet-Hoeffding upper bound.

Now suppose that $\ell \neq \ell_M$. Then it is clear that $C_k \neq C_M$ for both $k \in \{1, 2\}$. Indeed, fix $k \in \{1, 2\}$. Note that $\ell(\boldsymbol{x}) > \ell_M(\boldsymbol{x})$ for some $\boldsymbol{x} \in \mathbb{R}_+^d$. By the homogeneity of ℓ, there also exists an $\boldsymbol{x} \in \mathbb{R}_+^d$ such that $0 < \psi_k\{\ell(\boldsymbol{x})\} < \psi_k\{\ell_M(\boldsymbol{x})\}$. Therefore, $C_k(\boldsymbol{u}) = \psi_k \circ \ell\{\phi_k(\boldsymbol{u})\} < \psi_k \circ \ell_M\{\phi_k(\boldsymbol{u})\} = C_M(\boldsymbol{u})$ for $\boldsymbol{u} = \psi_k(\boldsymbol{x})$. Consequently, there exists at least one pair $i, j \in \{1, \dots, d\}$, $i < j$, such that the bivariate margin of C_k, given, for all $u_i, u_j \in [0, 1]$, by

$$C_k^{(ij)}(u_i, u_j) := C_k(1, \dots, 1, u_i, 1, \dots, 1, u_j, 1, \dots, 1)$$

is not the Fréchet-Hoeffding upper bound copula. Next note that for all $u_i, u_j \in [0, 1]$,

$$\begin{aligned} C_1^{(ij)}(u_i, u_j) &= \psi_1 \circ \ell^{(ij)} \{\phi_1(u_i), \phi_1(u_j)\} \\ &= \psi_2 \circ \ell^{(ij)} \{\phi_2(u_i), \phi_2(u_j)\} = C_2^{(ij)}(u_i, u_j), \end{aligned}$$

where $\ell^{(ij)}$ denotes the bivariate margin of ℓ, given, for all $x_i, x_j \in \mathbb{R}_+$, by

$$\ell^{(ij)}(x_i, x_j) = \ell(0, \dots, 0, x_i, 0, \dots, 0, x_j, 0, \dots, 0).$$

Therefore, $C_k^{(ij)}$, $k \in \{1, 2\}$ are bivariate Archimax. According to Equation (13) of Capéraà et al. (2000), they have the following Kendall's function for $w \in [0, 1]$,

$$K_k(w) = \tau_{\ell^{(ij)}} w + (1 - \tau_{\ell^{(ij)}}) K_{\psi_k}(w),$$

where $\tau_{\ell^{(ij)}}$ is the Kendall's tau of the extreme-value copula $C_{\ell^{(ij)}}$ and $K_{\psi_k}(w)$ is the Kendall's function of the bivariate Archimedean copula C_{ψ_k}. Since $\ell^{(ij)} \neq \ell_M$, we know that $\tau_{\ell^{(ij)}} < 1$ and thus that $K_{\psi_1}(w) = K_{\psi_2}(w)$. From Genest et al. (2011) and Genest and Rivest (1993), it follows that $C_{\psi_1} = C_{\psi_2}$. By the identifiability of Archimedean copulas, this yields the equality of ψ_1 and ψ_2 up to scaling (see for example Chapter 4 of Nelsen (2006)). □

The first part of the following lemma is an extension of Theorem 4.5.1 in Nelsen (2006) and has been shown by Hofert (2008) in the case where ψ is completely monotone. In the following, for any $\beta \in (0, 1]$, ψ_β is defined by $\psi_\beta(t) = \psi(t^\beta)$ for all $t \geq 0$, and ℓ_β denotes $\ell^\beta(x_1^{1/\beta}, \dots, x_d^{1/\beta})$ for all $\boldsymbol{x} \in \mathbb{R}_+^d$.

Lemma 3.3. *(i) Let ψ be a d-monotone Archimedean generator and $\beta \in (0, 1]$. Then ψ_β is a d-monotone Archimedean generator.*

(ii) Let ℓ be a d-variate stdf and $\beta \in (0, 1]$. Then ℓ_β is a d-variate stdf.

Proof. Proof of part (i). Clearly, ψ_β is a continuous and decreasing function such that $\psi_\beta(0) = 1$ and $\psi_\beta(x) \to 0$ as $x \to \infty$. Let ℓ_β be the logistic stdf given, for all $\boldsymbol{x} \in \mathbb{R}_+^d$ by $\ell_\beta(x_1, \ldots, x_d) = (x_1^{1/\beta} + \cdots + x_d^{1/\beta})^\beta$. The Archimax copula C_{ψ,ℓ_β}, is a bona-fide copula by Theorem 2.1 of Charpentier et al. (2014). However, it is easily seen that $C_{\psi,\ell_\beta} = C_{\psi_\beta}$, where C_{ψ_β} is the d-variate Archimedean copula with generator ψ_β. By Theorem 2.2 of McNeil and Nešlehová (2009), ψ_β must be d-monotone.

Proof of part (ii). Let ψ_β be the generator of the Gumbel copula given, for all $x \geq 0$, by $\psi_\beta(x) = e^{-x^\beta}$. Then ψ_β is a completely monotone Archimedean generator and $1 - \psi_\beta(1/x) \in \mathcal{R}_{-\beta}$. By Proposition 6.1 of Charpentier et al. (2014), the d-variate Archimax copula $C_{\psi_\beta,\ell}$ is in the maximum domain of attraction of the extreme-value copula with stdf ℓ_β. Consequently, ℓ_β is a d-variate stdf, as claimed. □

Now suppose that ψ is a d-monotone Archimedean generator and ℓ is an arbitrary d-variate stdf. By Lemma 3.3, ψ_β is a d-monotone Archimedean generator and ℓ_β is a d-variate stdf for some $\beta \in (0, 1]$. It is then easily seen that the Archimax copulas $C_{\psi_\beta,\ell}$ and C_{ψ,ℓ_β} coincide. Thus one cannot expect ℓ to be unique and ψ to be unique up to scaling. As stated below, however, under a mild regularity condition on ψ, power transformations of ψ and ℓ are the only possible sources of non-identifiability.

Lemma 3.4. *Suppose that $\ell_1 \neq \ell_M$ and $\ell_2 \neq \ell_M$ are arbitrary d-variate stdfs and ψ_1, ψ_2 are d-monotone Archimedean generators with the property that for $k \in \{1, 2\}$, $1-\psi_k(1/\cdot) \in \mathcal{R}_{-1/m_k}$, with $m_k \geq 1$. Assuming, without loss of generality, that $m_1 \leq m_2$, $C_{\psi_1,\ell_1} = C_{\psi_2,\ell_2}$ holds iff for all $\boldsymbol{x} \in \mathbb{R}_+^d$,*

$$\ell_1(x_1, \ldots, x_d) = \ell_2^{m_1/m_2}(x_1^{m_2/m_1}, \ldots, x_d^{m_2/m_1})$$

and there exists $c > 0$ such that, for all $t \geq 0$, $\psi_1(ct^{m_1/m_2}) = \psi_2(t)$.

Proof. Proposition 6.1 of Charpentier et al. (2014) implies that, for all $k \in \{1, 2\}$, that C_{ψ_k,ℓ_k} is in the maximum domain of attraction of the extreme-value copula with stdf given, for all $\boldsymbol{x} \in \mathbb{R}_+^d$, by $\ell_k^{1/m_k}(\boldsymbol{x}^{m_k})$. Because $C_{\psi_1,\ell_1} = C_{\psi_2,\ell_2}$ by assumption, this implies that for all $\boldsymbol{x} \in \mathbb{R}_+^d$, it holds that $\ell_1^{1/m_1}(\boldsymbol{x}^{m_1}) = \ell_2^{1/m_2}(\boldsymbol{x}^{m_2})$. Hence, for all $\boldsymbol{x} \in \mathbb{R}_+^d$,

$$\ell_1(x_1, \ldots, x_d) = \ell_2^{m_1/m_2}\left(x_1^{m_2/m_1}, \ldots, x_d^{m_2/m_1}\right) .$$

Thus, for all $\boldsymbol{u} \in [0, 1]^d$,

$$C_{\psi_1,\ell_1}(\boldsymbol{u}) = \psi_1 \circ \ell_2^{m_1/m_2}\left[\left\{\psi_1^{-1}(u_1)\right\}^{m_2/m_1}, \ldots, \left\{\psi_1^{-1}(u_d)\right\}^{m_2/m_1}\right] .$$

Now set $\psi_1^\star(t) = \psi_1\left(t^{m_1/m_2}\right)$ for $t \in \mathbb{R}_+$ and note that $\psi_1^\star$ is a d-montone Archimedean generator by Lemma 3.3. Therefore, $C_{\psi_1,\ell_1} = C_{\psi_1^\star,\ell_2} = C_{\psi_2,\ell_2}$. Given that $\ell_2 \neq \ell_M$ by assumption, the rest of the claim follows from Lemma 3.2. □

Lemma 3.4 allows us to formulate the following main result of this section that delineates the conditions under which an Archimax copula model is identifiable assuming that the Archimedean generator belongs to a parametric family. Its proof is a direct consequence of Lemma 3.4.

Proposition 3.1. *Let $\mathcal{C}_\Psi$ be a class of d-variate Archimax copulas whose stdfs are arbitrary with $\ell \neq \ell_M$ and whose Archimedean generators belong to $\Psi = \{\psi_\theta, \theta \in \mathcal{O}\}$, $\mathcal{O} \subset \mathbb{R}^p$. Assume also that the following conditions hold:*

(i) for all $\theta \in \mathcal{O}$, $1 - \psi_\theta(1/\cdot) \in \mathcal{R}_{-1/m_\theta}$, with $m_\theta \geq 1$;

(ii) for all $\theta \in \mathcal{O}$, $c > 0$, and $\beta > 0$, the function given, for all $t \geq 0$ by $\psi_\theta(ct^\beta)$ is an element of Ψ if and only if $c = \beta = 1$.

Then for any $C_{\psi_\theta,\ell}, C_{\psi_{\theta'},\ell'} \in \mathcal{C}_\Psi$, $C_{\psi_\theta,\ell} = C_{\psi_{\theta'},\ell'}$ holds iff $\ell = \ell'$ and $\theta = \theta'$.

Condition (i) in Proposition 3.1 returns as Condition 3.1 in Section 4.1, where it is discussed in detail. As shown by Charpentier and Segers (2009), it holds for many Archimedean families, including those in Table 4.1 of Nelsen (2006). Condition (ii) is satisfied by most commonly used one-parameter families of Archimedean generators, e.g., the Ali–Mikhail–Haq, Clayton, and Frank models. The only exceptions we could find are Families 4.2.2, 4.2.4 (Gumbel), 4.2.12, and 4.2.18 in Nelsen (2006), and the outer power family $\phi_{1,\beta}$ from Theorem 4.5.1 therein. Lack of identifiability is not a concern for these models, however, because through Lemma 3.4, θ can be absorbed into the stdf so that the generator ψ of the resulting Archimax model is fixed. For example, for the Gumbel generator given by $\psi_\theta(x) = e^{-x^{1/\theta}}$, and an arbitrary d-variate stdf ℓ, the Archimax copula $C_{\psi_\theta,\ell}$ coincides with the Archimax copula C_{ψ_1,ℓ_θ}, where the Archimedean generator $\psi_1(x) = e^{-x}$ no longer contains any parameters, and $\ell_\theta(\boldsymbol{x}) = \ell^{1/\theta}(\boldsymbol{x}^\theta)$.

3.2 Smoothness

The result in Berghaus et al. (2017) requires smoothness assumptions, namely Conditions 2.1 and 2.2 in the previous chapter. These are the same assumptions that appear in Segers (2012). We verify that these conditions indeed hold for Archimax copulas under suitable assumptions on the generator and the stdf, and this is nontrivial. To start, these said assumptions on ψ and ℓ are stated and discussed.

Condition 3.1. *For $d \geq 2$, ψ is a d-monotone Archimedean generator and $1 - \psi(1/x) \in \mathcal{R}_{-1/m}$ for some $m \geq 1$.*

Condition 3.1, which is equivalent to $\phi(1 - 1/x) \in \mathcal{R}_{-m}$, is very general and satisfied by virtually all d-monotone Archimedean generators as seen in Charpentier and Segers (2009); Larsson and Nešlehová (2011). This is because it holds whenever $1/R$ with R as in (2.11), is in the domain of attraction of the Fréchet (Φ_α), Gumbel (Λ) or Weibull (Ψ_α) distributions for some $\alpha > 0$, in notation $1/R \in \mathcal{M}(\Phi_\alpha)$, $1/R \in \mathcal{M}(\Lambda)$ or $1/R \in \mathcal{M}(\Psi_\alpha)$. Moreover, Condition 3.1 with $m = 1$ further holds as soon as $\mathrm{E}(1/R^{1+\epsilon}) < \infty$ for some $\epsilon > 0$; see Proposition 2 in Belzile and Nešlehová (2017).

Condition 3.2. *For $d \geq 2$, ψ is a d-monotone Archimedean generator that satisfies either*

(a) $\psi \in \mathcal{R}_{-s}$ *for* $s > 0$*;*

(b) $Y \in \mathcal{M}(\Lambda)$*, where* Y *has distribution function* $1 - \psi$*;*

(c) $\phi(0) < \infty$ *and* $\psi(x_\psi - 1/x) \in \mathcal{R}_{-\alpha-d+1}$ *for* $\alpha > 0$.

Most Archimedean generators satisfy Condition 3.2. As shown by Larsson and Nešlehová (2011), Condition 3.2 (a) holds whenever R in (2.11) is such that $R \in \mathcal{M}(\Phi_s)$ and is further equivalent to $\phi(1/x) \in \mathcal{R}_{1/s}$. Condition 3.2 (b) is equivalent to $1/\psi$ being Γ-varying which is in turn equivalent to $\phi(1/x)$ being Π-varying, as defined and proved, e.g., in Section 0.4.3 in Resnick (1987). It is further shown by Larsson and Nešlehová (2011) that Condition 3.2 (b) holds whenever $R \in \mathcal{M}(\Lambda)$. Finally, Condition 3.2 (c) is equivalent to $R \in \mathcal{M}(\Psi_\alpha)$ and further to $\{\phi(0) - \phi(1/x)\} \in \mathcal{R}_{-1/(\alpha+d-1)}$.

Condition 3.3. *For $d \geq 2$, ℓ is a d-variate stdf that is twice continuously differentiable and for which there exists $M > 0$ such that for any $i, j \in \{1, \ldots, d\}$ with $i \neq j$, and for any $\boldsymbol{x} \in (0, \infty)^d$,*

$$-\frac{\partial^2}{\partial x_i \partial x_j} \ell(x_1, \ldots, x_d) \equiv -\ddot{\ell}_{ij}(x_1, \ldots, x_d) \leq M \left(\frac{1}{x_i} \wedge \frac{1}{x_j} \right).$$

Condition 3.3 extends Condition 5.2 in Segers (2012) to the case $d > 2$. The following example demonstrates that it is satisfied by the logistic stdf.

Example 3.1. *The logistic stdf is given for any $\boldsymbol{x} \in \mathbb{R}_+^d$ and $\theta \geq 1$ by $\ell_\theta(x_1, \ldots, x_d) = (x_1^\theta + \ldots + x_d^\theta)^{1/\theta}$. It is easily seen that for any $\boldsymbol{x} \in \mathbb{R}_+^d$,*

$$-\ddot{\ell}_{ij}(\boldsymbol{x}) = (\theta - 1) x_i^{\theta-1} x_j^{\theta-1} \left(x_1^\theta + \cdots + x_d^\theta\right)^{1/\theta - 2} \leq (\theta - 1) \left(\frac{1}{x_i} \wedge \frac{1}{x_j} \right).$$

Proposition 3.2 below is the main result of this section, as it delineates the assumptions under which Conditions 2.1 and 2.2 hold.

Proposition 3.2. *Suppose that $C_{\psi,\ell}$ is a d-variate Archimax copula with Archimedean generator ψ that is q-monotone for some $q \geq 0$ and such that ψ'' exists and is continuous on $(0, \infty)$. Further assume that Conditions 3.1 and 3.3, and that either Condition 3.2 (a) is satisfied or Condition 3.2 (b) is satisfied with the additional requirement that $-\log \psi$ is concave on $(0, x_\psi)$. Then Conditions 2.1 and 2.2 are met.*

Remark 3.1. *Proposition 3.2 also shows that Condition (4.1) in Segers (2012) holds for an Archimedean copula C_ψ if ψ is q-monotone for some $q \geq 3$, ψ'' exists and is continuous on $(0, \infty)$, Condition 3.1 holds, and either Condition 3.2 (a) is satisfied or Condition 3.2 (b) is satisfied with the additional requirement that $-\log(\psi)$ is concave.*

The proof of proposition 3.2 requires many auxiliary results that are presented in Section 3.2.1 below. The result is then proved in two parts, formulated as Propositions 3.3 and 3.4 in Section 3.2.2.

3.2.1 Auxiliary results

Before getting to the main results, some auxiliary results are needed. Let C be a d-dimensional Archimax copula $C_{\psi,\ell}$. With the notation $\phi(\boldsymbol{u}) = \{\phi(u_1), \ldots, \phi(u_d)\}$, the partial derivatives of C can be computed for each $i, j \in \{1, \ldots, d\}$, $i \neq j$, as

$$\dot{C}_i(\boldsymbol{u}) = \psi' [\ell\{\phi(\boldsymbol{u})\}] \, \dot{\ell}_i\{\phi(\boldsymbol{u})\}\phi'(u_i), \tag{3.1}$$

$$\begin{aligned} \ddot{C}_{ij}(\boldsymbol{u}) = \Big(&\psi'' [\ell\{\phi(\boldsymbol{u})\}] \, \dot{\ell}_i\{\phi(\boldsymbol{u})\}\dot{\ell}_j\{\phi(\boldsymbol{u})\} + \psi' [\ell\{\phi(\boldsymbol{u})\}] \, \ddot{\ell}_{ij}\{\phi(\boldsymbol{u})\}\Big) \\ &\times \phi'(u_i)\phi'(u_j), \end{aligned} \tag{3.2}$$

$$\begin{aligned} \ddot{C}_{ii}(\boldsymbol{u}) = \Big(&\psi'' [\ell\{\phi(\boldsymbol{u})\}] \, [\dot{\ell}_i\{\phi(\boldsymbol{u})\}]^2 + \psi' [\ell\{\phi(\boldsymbol{u})\}] \, \ddot{\ell}_{ii}\{\phi(\boldsymbol{u})\}\Big) \\ &\times \{\phi'(u_i)\}^2 + \psi' [\ell\{\phi(\boldsymbol{u})\}] \, \dot{\ell}_i\{\phi(\boldsymbol{u})\}\phi''(u_i) \; . \end{aligned} \tag{3.3}$$

Lemma 3.5. *Let ℓ be a $d-$variate stdf whose first order partial derivatives exist on $\mathbb{R}^d_+$. Then, for any $i \in \{1, \ldots, d\}$ and $\boldsymbol{x} \in \mathbb{R}^d_+$, $0 \leq \dot{\ell}_i(\boldsymbol{x}) \leq 1$.*

Proof. Both inequalities can be derived from the properties (a)–(c) in Theorem 2.6. Fix $i \in \{1, \ldots, d\}$ and $\boldsymbol{x} \in \mathbb{R}^d_+$. Since ℓ is fully d-max decreasing, it is increasing in each argument. This yields the first inequality. To show the second inequality, note that properties (a) and (b) imply $\ell(0, \ldots, 0, x_i, 0, \ldots, 0) = x_i$, and hence $\dot{\ell}_i(0, \ldots, 0, x_i, 0, \ldots, 0) = 1$. From property (c), it also follows that $\dot{\ell}_i$ is non-increasing in the j-th argument for all $j \neq i$. Therefore $\dot{\ell}_i(\boldsymbol{x}) \leq \dot{\ell}_i(0, \ldots, 0, x_i, 0, \ldots, 0) = 1$. □

Lemma 3.6. *Let ψ be a d-monotone Archimedean generator for some $d \geq 2$ such that ψ' exists and is continuous on $(0, \infty)$ when $d = 2$. Assume that Conditions 3.1 and 3.2 hold and let $x_\psi = \inf\{x \in [0, \infty) : \psi(x) = 0\}$. Then the function given for any $x \in (0, x_\psi)$ by $f(x) = \psi(x)\{1 - \psi(x)\}/\{-x\psi'(x)\}$ is continuous on $(0, x_\psi)$ and has finite limits at 0 and x_ψ.*

Proof. Given that the continuity of f is immediate, it suffices to show that its limits at 0 and x_ψ are finite. Because Condition 3.1 holds,

$$\lim_{x\to 0} f(x) = \lim_{x\to\infty} \frac{\psi(1/x)\{1-\psi(1/x)\}}{(1/x)\{-\psi'(1/x)\}} = m ,$$

where the last equality follows from Equation (12) of Larsson and Nešlehová (2011). Turning to the limit of f at x_ψ, three cases have to be distinguished.

Assume first that Condition 3.2 (a) holds. In this case, $x_\psi = \infty$ and Equation (7) of Larsson and Nešlehová (2011) implies $\lim_{x\to\infty} f(x) = 1/s$. Next, assume that Condition 3.2 (b) holds. Because the function given for all $x \in (0, x_\psi)$ by $\psi(x)/\{-\psi'(x)\}$ is an auxiliary function by the calculations in the proof of Theorem 1 (c) on p. 213 of Larsson and Nešlehová (2011), $\lim_{x\to x_\psi} f(x) = 0$ by Lemma 3.10.1 of Bingham et al. (1989). Finally, assuming Condition 3.2 (c), $x_\psi < \infty$ and

$$\begin{aligned}
\lim_{x\to x_\psi} f(x) &= \lim_{x\to\infty} \frac{\{1-\psi(x_\psi - 1/x)\}\psi(x_\psi - 1/x)}{-\psi'(x_\psi - 1/x)(x_\psi - 1/x)} \\
&= \lim_{x\to\infty} \frac{x\psi(x_\psi - 1/x)}{-\psi'(x_\psi - 1/x)} \, \frac{(1/x)\{1-\psi(x_\psi - 1/x)\}}{x_\psi - 1/x} = 0 ,
\end{aligned}$$

since the first ratio in the last expression tends to $1/(\alpha + d - 1)$ thanks to Condition 3.2 (c) and the proof of Theorem 1 (b) on p. 211 of Larsson and Nešlehová (2011). □

Lemma 3.7. *Let ψ be a d-monotone Archimedean generator for some $d \ge 3$ such that ψ'' exists and is continuous on $(0,\infty)$. Assume that Conditions 3.1 and 3.2 hold and let $x_\psi = \inf\{x \in [0,\infty) : \psi(x) = 0\}$. Then the function given for any $x \in (0, x_\psi)$ by $f(x) = \psi(x)\{1-\psi(x)\}\psi''(x)/\{\psi'(x)\}^2$ is continuous on $(0, x_\psi)$ and has finite limits at 0 and x_ψ.*

Proof. As in the proof of Lemma 3.6, the continuity of f is immediate and hence it suffices to show that its limits at 0 and x_ψ are finite. From Condition 3.1 and Equation (12) of Larsson and Nešlehová (2011),

$$\lim_{x\to 0} f(x) = \lim_{x\to\infty} \frac{\psi(1/x)\{1-\psi(1/x)\}(1/x)^2\psi''(1/x)}{\{-(1/x)\psi'(1/x)\}^2} = m - 1 ,$$

Turning to the limit of f at x_ψ, three cases have to be distinguished.

Assume first that Condition 3.2 (a) holds. In this case, $x_\psi = \infty$ and Equation (7) of Larsson and Nešlehová (2011) implies $\lim_{x\to\infty} f(x) = (s+1)/s$.

Next, assume that Condition 3.2 (b) holds. By the calculations in the proof of Theorem 1 (c) on p. 213 of Larsson and Nešlehová (2011), the functions given for all $x \in (0, x_\psi)$ by $a_1^*(x) = \psi(x)/\{-\psi'(x)\}$ and $a_2^*(x) = -\psi'(x)/\psi''(x)$ are auxiliary functions that are asymptotically equivalent to the auxiliary function a of ψ. Consequently, $a_1^*(x)/a_2^*(x) \to 1$ as $x \to x_\psi$ so that $\lim_{x\to x_\psi} f(x) = 1$.

Finally, assuming Condition 3.2 (c), $x_\psi < \infty$ and

$$\lim_{x \to x_\psi} f(x) = \lim_{x \to \infty} \frac{\psi(x_\psi - 1/x)(1/x)^2 \psi''(x_\psi - 1/x)}{\{-(1/x)\psi'(x_\psi - 1/x)\}^2} \{1 - \psi(x_\psi - 1/x)\}$$
$$= \frac{\alpha + d - 2}{\alpha + d - 1} \,,$$

where the last equality follows from the calculations on p. 211 in the proof of Theorem 1 (b) of Larsson and Nešlehová (2011). □

3.2.2 Proof of Proposition 3.2

Proposition 3.2 is an immediate consequence of the following two propositions.

Proposition 3.3. *Let $C = C_{\psi,\ell}$ be a d-variate Archimax copula such that ψ' exists and is continuous on $(0, \infty)$ when $d = 2$, and the first order partial derivatives of ℓ exist and are continuous on $\mathbb{R}_+^d$. Then Condition 2.1 holds.*

Proof. Fix $j \in \{1, \ldots, d\}$, $\boldsymbol{u} \in V_{d,j}$, set $\boldsymbol{x} = \phi(\boldsymbol{u})$ and using (3.1) write

$$\dot{C}_j\{\psi(\boldsymbol{x})\} = \frac{\psi'\{\ell(\boldsymbol{x})\}\dot{\ell}_j(\boldsymbol{x})}{\psi'(x_j)}.$$

Because $\psi' > 0$ on $(0, x_\psi)$, and $\ell(\boldsymbol{x}) \geq x_j > 0$ on $V_{d,j}$, the assumptions imply that $\dot{C}_j$ is continuous on $(0, 1]^d \cap V_{d,j}$. If $u_i \to 0$ for at least one $i \neq j$, $x_i \to \phi(0)$ and $\ell(\boldsymbol{x}) \to \ell(x_1, \ldots, x_{i-1}, \phi(0), x_{i+1}, \ldots, x_d) \geq \phi(0)$. By Lemma 1 of Williamson (1956), $\psi'(x) \to 0$ as $x \to \phi(0)$ and if $\phi(0) < \infty$, $\psi'(x) = 0$ for $x \geq \phi(0)$. Consequently, as $x_i \to \phi(0)$, $\dot{C}_j\{\psi(\boldsymbol{x})\} \to 0$. □

Proposition 3.4. *Let $C = C_{\psi,\ell}$ be a d-variate Archimax copula such that ψ is k-monotone for some $k \geq 3$ and ψ'' exists and is continuous on $(0, \infty)$. If Conditions 3.1, 3.2 (a) and 3.3 hold, or if $-\log(\psi)$ is concave and Conditions 3.1, 3.2 (b) and 3.3 hold, then Condition 2.2 is satisfied.*

Proof. For any $\boldsymbol{u} \in [0, 1]^d$, set $\boldsymbol{x} = \phi(\boldsymbol{u})$ and for any $i, j \in \{1, \ldots, d\}$, introduce the following terms:

$$T_{ij,1}(\boldsymbol{x}) = \frac{\psi''\{\ell(\boldsymbol{x})\}}{\psi'(x_i)\psi'(x_j)}, \quad T_{ij,2}(\boldsymbol{x}) = \frac{-M\psi'\{\ell(\boldsymbol{x})\}}{(x_i \vee x_j)\psi'(x_i)\psi'(x_j)},$$
$$T_{ii,3}(\boldsymbol{x}) = \frac{\psi'\{\ell(\boldsymbol{x})\}\psi''(x_i)}{\{\psi'(x_i)\}^3}.$$

By the d-monotonicity of ψ, observe first that for $k \in \{1, 2, 3\}$, $T_{ij,k} \geq 0$. Now let $x_\psi = \inf\{x \in [0, \infty) : \psi(x) = 0\}$. From (3.2), (3.3), Lemma 3.5, and Condition 3.3 it follows that for any $\boldsymbol{x} \in (0, x_\psi)^d$,

$$|\ddot{C}_{ij}\{\psi(\boldsymbol{x})\}| \leq T_{ij,1}(\boldsymbol{x}) + T_{ij,2}(\boldsymbol{x}), \quad |\ddot{C}_{ii}\{\psi(\boldsymbol{x})\}| \leq T_{ii,1}(\boldsymbol{x}) + T_{ii,2}(\boldsymbol{x}) + T_{ii,3}(\boldsymbol{x}).$$

Next, note that for any $i \neq j$, $\ddot{C}_{ij}$ and $\ddot{C}_{ii}$ are continuous on $(0,1]^d \cap V_{d,i} \cap V_{d,j}$. The d-monotonicity of ψ and Lemma 1 of Williamson (1956) implies that for $k \in \{1,2\}$, $\psi^{(k)}(x) \to 0$ as $x \to x_\psi$ and if $x_\psi < \infty$, $\psi^{(k)}(x) = 0$ for $x \geq x_\psi$. Consequently, for each $k \in \{1,2\}$, $T_{ij,k}(\boldsymbol{x}) \to 0$ as $x_r \to x_\psi$ for at least one $r \notin \{i,j\}$ and that for each $k \in \{1,2,3\}$, $T_{ii,k}(\boldsymbol{x}) \to 0$ as $x_r \to x_\psi$ for at least one $r \neq i$. This in turn implies that $\ddot{C}_{ij}\{\psi(\boldsymbol{x})\} \to 0$ and $\ddot{C}_{ii}\{\psi(\boldsymbol{x})\} \to 0$ as $x_r \to x_\psi$ for at least one r in $\{1,\ldots,d\} \setminus \{i,j\}$ and $\{1,\ldots,d\} \setminus \{i\}$, respectively. Hence for $i \neq j$, $\ddot{C}_{ij}$ and $\ddot{C}_{ii}$ are continuous on $V_{d,i} \cap V_{d,j}$.

Now introduce the functions given, for any $z_1, z_2 \in (0, x_\psi)$, by

$$\tilde{T}_1(z_1,z_2) = \frac{\psi''\{z_1 \vee z_2\}}{\psi'(z_1)\psi'(z_2)}, \quad \tilde{T}_2(z_1,z_2) = \frac{-M\psi'\{z_1 \vee z_2\}}{(z_1 \vee z_2)\psi'(z_1)\psi'(z_2)},$$
$$\tilde{T}_3(z_1) = \frac{\psi'\{z_1\}\psi''(z_1)}{\{\psi'(z_1)\}^3}.$$

Note first that for $k \in \{1,2,3\}$, $\tilde{T}_k \geq 0$ on its domain. Because $(-1)^q\psi^{(q)}$ is nonincreasing on $[0,\infty)$ for $q \in \{1,2\}$ and $\ell(\boldsymbol{x}) \geq x_1 \vee \cdots \vee x_d$ for any $\boldsymbol{x} \in \mathbb{R}_+^d$, one has that for any $i \neq j$ and any $\boldsymbol{x} \in \{\phi(\boldsymbol{u}), \boldsymbol{u} \in V_{d,i} \cap V_{d,j}\}$ and $\boldsymbol{x} \in \{\phi(\boldsymbol{u}), \boldsymbol{u} \in V_{d,i}\}$,

$$|\ddot{C}_{ij}\{\psi(\boldsymbol{x})\}| \leq \tilde{T}_1(x_i,x_j) + \tilde{T}_2(x_i,x_j)$$
$$\text{and} \quad |\ddot{C}_{ii}\{\psi(\boldsymbol{x})\}| \leq \tilde{T}_1(x_i,x_i) + \tilde{T}_2(x_i,x_i) + \tilde{T}_3(x_i),$$

respectively. Note that for $k \in \{1,2\}$, the term $\tilde{T}_k$ is symmetric. To show the inequality

$$|\ddot{C}_{ij}(\boldsymbol{u})| \leq K \min\left\{\frac{1}{u_i(1-u_i)}, \frac{1}{u_j(1-u_j)}\right\}$$

it thus suffices to prove that for $k \in \{1,2\}$, the function given for all $z_1, z_2 \in (0,x_\psi)$ by $\psi(z_1)\{1-\psi(z_1)\}\tilde{T}_k(z_1,z_2)$ is bounded on $(0,x_\psi)^2$, and further that the function given for all $z_1 \in (0,x_\psi)$ by $\psi(z_1)\{1-\psi(z_1)\}\tilde{T}_3(z_1)$ is bounded on $(0,x_\psi)$. First observe that because $-\psi'$ is nonincreasing,

$$\psi(z_1)\{1-\psi(z_1)\}\tilde{T}_2(z_1,z_2) \leq \frac{M\psi(z_1)\{1-\psi(z_1)\}}{-z_1\psi'(z_1)},$$
$$\psi(z_1)\{1-\psi(z_1)\}\tilde{T}_3(z_1) \leq \frac{\psi''(z_1)\psi(z_1)\{1-\psi(z_1)\}}{\{\psi'(z_1)\}^2}\,.$$

The function on the right-hand side in the first and second inequality is bounded on $(0,x_\psi)$ by Lemma 3.6 and Lemma 3.7, respectively.

It remains to consider the function $\tilde{T}_1$. For all $z_1, z_2 \in (0,x_\psi)$, denote $h(z_1,z_2) = \psi(z_1)\{1-\psi(z_1)\}\tilde{T}_1(z_1,z_2)$. First note that because $-\psi'$ is decreasing on $(0,x_\psi)$,

$$h(z_1,z_2) \leq \frac{\psi(z_1)\{1-\psi(z_1)\}}{-z_1\psi'(z_1)} \frac{(z_1 \vee z_2)\psi''(z_1 \vee z_2)}{-\psi'(z_1 \vee z_2)} = f(z_1)g(z_1 \vee z_2)\,, \tag{3.4}$$

in terms of $f(x) = \psi(x)\{1 - \psi(x)\}/\{-x\psi'(x)\}$ and $g(x) = x\psi''(x)/\{-\psi'(x)\}$. Now f is bounded on $(0, x_\psi)$ by Lemma 3.6. Furthermore, g is continuous and because Condition 3.1 holds, it satisfies

$$\lim_{x\to 0} g(x) = \lim_{x\to\infty} \frac{(1/x)^2\psi''(1/x)}{(1/x)\left\{-\psi'(1/x)\right\}} = 1 - 1/m \ ,$$

where the last equality follows from Equation (12) of Larsson and Nešlehová (2011). Therefore, h is bounded on $(0, \kappa]^2$ for any $\kappa < x_\psi$. To conclude that h is bounded on the entire set $(0, x_\psi)^2$, two cases have to be distinguished. First, assume that Condition 3.2 (a) holds. In this case, $x_\psi = \infty$ and Equation (7) of Larsson and Nešlehová (2011) implies $\lim_{x\to\infty} g(x) = s + 1$ and hence the upper bound in (3.4) is bounded on $(0, x_\psi)^2$. Next, assume that Condition 3.2 (b) holds, and that $-\log(\psi)$ is concave. In this case, the upper bound in (3.4) is too crude because $g(x) \to \infty$ as $x \to x_\psi$. Instead observe that, because ψ is decreasing,

$$\begin{aligned} h(z_1, z_2) &= \frac{\psi(z_1 \vee z_2)\psi''(z_1 \vee z_2)}{\{\psi'(z_1 \vee z_2)\}^2} \frac{\psi(z_1)}{\psi(z_1 \vee z_2)} \frac{\psi'(z_1 \vee z_2)}{\psi'(z_1 \wedge z_2)} \\ &\le \frac{\psi(z_1 \vee z_2)\psi''(z_1 \vee z_2)}{\{\psi'(z_1 \vee z_2)\}^2} \frac{a_1^*(z_1 \wedge z_2)}{a_1^*(z_1 \vee z_2)}, \end{aligned} \tag{3.5}$$

where for any $x \in (0, x_\psi)$, $a_1^*(x) = \psi(x)/\{-\psi'(x)\}$. From the proof of Lemma 3.7, $\psi(x)\psi''(x)/\{\psi'(x)\}^2 \to 1$ as $x \to x_\psi$. Furthermore, because $-\log(\psi)$ is concave, a_1^* is increasing and hence the upper bound in (3.5) is bounded on $(0, x_\psi)^2 \setminus (0, \kappa]^2$ for any $\kappa \in (0, x_\psi)$. Put together, h is bounded on $(0, x_\psi)^2$. □

Chapter 4

Estimating ℓ when ψ is known

In this chapter, we introduce two nonparametric estimators of the stdf ℓ of an Archimax copula $C_{\psi,\ell}$ under the assumption that the Archimedean generator ψ is known. As stated in Chapter 3, ℓ is identifiable under this assumption. Recall that ℓ is uniquely determined by the corresponding Pickands dependence function A, and hence it suffices to estimate the latter. To see how to proceed, consider a random vector $\boldsymbol{U}$ with distribution $C_{\psi,A}$ given by (2.9). For any $\boldsymbol{w}$ in the unit simplex Δ_d, let

$$\xi(\boldsymbol{w}) = \min\{\phi(U_1)/w_1, \ldots, \phi(U_d)/w_d\}$$

with $\phi(U_j)/w_j = \infty$ when $w_j = 0$ for some $j \in \{1, \ldots, d\}$. Then

$$\Pr\{\xi(\boldsymbol{w}) > x\} = C_{\psi,A}\{\psi(x\boldsymbol{w})\} = \psi\{xA(\boldsymbol{w})\}.$$

If $\psi(x) = e^{-x}$, $\xi(\boldsymbol{w})$ is exponential with rate $A(\boldsymbol{w})$. This leads to Pickands and Capéraà–Fougères–Genest (CFG) type estimators of A; these estimators are investigated, e.g., in Pickands (1981); Capéraà et al. (1997); Zhang et al. (2008); Genest and Segers (2009); Gudendorf and Segers (2011).

Now let Z denote a random variable with survival function ψ, i.e., for all $x \geq 0$, $\Pr(Z > x) = \psi(x)$. Then for any $\boldsymbol{w} \in \Delta_d$, $\xi(\boldsymbol{w})$ has the same distribution as $Z/A(\boldsymbol{w})$. One finds in particular that

$$\mathrm{E}\{\xi(\boldsymbol{w})\} = \mathrm{E}(Z)/A(\boldsymbol{w}), \quad \mathrm{E}[\log\{\xi(\boldsymbol{w})\}] = \mathrm{E}(\log Z) - \log\{A(\boldsymbol{w})\}. \tag{4.1}$$

When ψ is known, so are $\mathrm{E}(Z)$ and $\mathrm{E}(\log Z)$. Provided the latter are finite, (4.1) leads to the Pickands and CFG-type estimators of A, as explained next.

Let $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n$ be a random sample from a d-variate distribution H with continuous margins $F_1, \ldots, F_d$ and an Archimax copula $C_{\psi,A}$ with known ψ and unknown A. When the margins are unknown, a sample from $C_{\psi,A}$ is unavailable, but as in Genest and Segers (2009) and Gudendorf and Segers (2012), one can base inference on normalized ranks given, for all $i \in \{1, \ldots, n\}$ and $j \in \{1, \ldots, d\}$ by $\hat{U}_{ij} = n\,F_{nj}(X_{ij})/(n+1)$, where for any $j \in \{1, \ldots, d\}$, F_{nj} is the empirical distribution function of $X_{1j}, \ldots, X_{nj}$, as defined in Equation (2.13). Now, for every $\boldsymbol{w} \in \Delta_d$ and $i \in \{1, \ldots, n\}$, let

$$\hat{\xi}_i(\boldsymbol{w}) = \min\{\phi(\hat{U}_{i1})/w_1, \ldots, \phi(\hat{U}_{id})/w_d\}$$

again with the convention that $\phi(\hat{U}_{ij})/w_j = \infty$ when $w_j = 0$. However, note that for any $\boldsymbol{w} \in \Delta_d$, $w_j > 0$ for at least one j, so that $\hat{\xi}_i(\boldsymbol{w})$ is finite for every $i \in \{1, \ldots, n\}$. Then, provided that $\mathrm{E}(Z)$ exists, the Pickands-type estimator A_n^{P} is defined, for any $\boldsymbol{w} \in \Delta_d$, by

$$A_n^{\mathrm{P}}(\boldsymbol{w}) = n\,\mathrm{E}(Z) \Big/ \sum_{i=1}^{n} \hat{\xi}_i(\boldsymbol{w}). \tag{4.2}$$

Similarly, if $\mathrm{E}(\log Z)$ exists, the CFG-type estimator A_n^{CFG} is defined through

$$\log A_n^{\mathrm{CFG}}(\boldsymbol{w}) = \mathrm{E}\log Z - \frac{1}{n}\sum_{i=1}^{n} \log \hat{\xi}_i(\boldsymbol{w}). \tag{4.3}$$

If $\psi(x) = e^{-x}$, then $\mathrm{E}(Z) = 1$ and $\mathrm{E}(\log Z) = -\gamma$, where γ is the Euler–Mascheroni constant, and A_n^{P} and A_n^{CFG} reduce to the rank-based Pickands and CFG estimators studied by Genest and Segers (2009) in dimension $d = 2$ and extended to higher dimensions by Gudendorf and Segers (2012).

In general, A_n^{P} and A_n^{CFG} are not Pickands dependence functions. In order to enforce the endpoint constraints $A(\boldsymbol{e}_j) = 1$ for $j \in \{1, \ldots, d\}$, introduce

$$\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} \phi\left(\frac{i}{n+1}\right), \quad \hat{\nu} = \frac{1}{n}\sum_{i=1}^{n} \log\phi\left(\frac{i}{n+1}\right).$$

The endpoint-corrected Pickands and CFG-type estimators now arise by replacing $\mathrm{E}(Z)$ by $\hat{\mu}$ in (4.2) and $\mathrm{E}(\log Z)$ by $\hat{\nu}$ in (4.3), respectively, viz.

$$A_{n,c}^{\mathrm{P}}(\boldsymbol{w}) = n\hat{\mu} \Big/ \sum_{i=1}^{n} \hat{\xi}_i(\boldsymbol{w}), \quad \log A_{n,c}^{\mathrm{CFG}}(\boldsymbol{w}) = \hat{\nu} - \frac{1}{n}\sum_{i=1}^{n} \log \hat{\xi}_i(\boldsymbol{w}). \tag{4.4}$$

These corrected versions avoid the generally cumbersome computation of $\mathrm{E}(Z)$ or $\mathrm{E}(\log Z)$. In addition, the following holds, owing to the fact that $\hat{\mu} = \sum_{i=1}^{n} \phi(\hat{U}_{ij})/n$ and $\hat{\nu} = \sum_{i=1}^{n} \log\phi(\hat{U}_{ij})/n$ almost surely for all $j \in \{1, \ldots, d\}$.

Proposition 4.1. *For $j \in \{1, \ldots, d\}$, $A_{n,c}^{\mathrm{P}}(\boldsymbol{e}_j) = 1$ and $A_{n,c}^{\mathrm{CFG}}(\boldsymbol{e}_j) = 1$ almost surely. Moreover, $A_{n,c}^{\mathrm{P}}(\boldsymbol{w}) \geq \max(w_1, \ldots, w_d)$ and $A_{n,c}^{\mathrm{CFG}}(\boldsymbol{w}) \geq \max(w_1, \ldots, w_d)$ almost surely for all $\boldsymbol{w} \in \Delta_d$.*

Note that when $d = 2$ and $\psi(x) = e^{-x}$, $A_{n,c}^{\mathrm{P}}$ is the corrected rank-based Pickands estimator from Genest and Segers (2009) with end-point correction as by Hall and Tajvidi (2000).

4.1 Asymptotic behavior

In this section, we investigate the asymptotic behavior of the Pickands and CFG-type estimators under the assumption that ψ is known. This section elicits the limiting behavior of the processes

$$\mathbb{A}_n^{\mathrm{P}} = \sqrt{n}\left(A_n^{\mathrm{P}} - A\right) \quad \text{and} \quad \mathbb{A}_n^{\mathrm{CFG}} = \sqrt{n}\left(A_n^{\mathrm{CFG}} - A\right). \tag{4.5}$$

The main ingredients of the proof are then made explicit in Section 4.2.

The following Lemma explains that under Conditions 3.1 and 3.2 studied in Section 3.2 of the previous chapter, the Pickands and CFG-type estimators are indeed well-defined and have the same limiting behavior as their end-point corrected versions.

Lemma 4.1. *(i) Suppose that ψ is differentiable on $(0, \infty)$ and satisfies either Condition 3.2 (a) with $s > 1$, (b) or (c). Then $\mathrm{E}(Z) < \infty$ and $\hat{\mu} \to \mathrm{E}(Z)$ as $n \to \infty$.*

(ii) Suppose that ψ is differentiable on $(0, \infty)$ and satisfies Conditions 3.1 and 3.2. Then $\mathrm{E}(\log Z) < \infty$ and $\hat{\nu} \to \mathrm{E}(\log Z)$ as $n \to \infty$.

Proof. For part (i), note that Condition 3.2 (a) with $s > 1$ is equivalent to $Z \in \mathcal{M}(\Phi_s)$ with $s > 1$. Similarly, Condition 3.2 (b) is equivalent to $Z \in \mathcal{M}(\Lambda)$, and Condition 3.2 (c) implies that Z is bounded from above. In either case, $\mathrm{E}(Z) < \infty$, see, e.g., Chapter 3 of Embrechts et al. (1997). Before showing that $\hat{\mu} \to \mathrm{E}(Z)$ as $n \to \infty$, note that for any positive random variable with finite expectation and a differentiable survival function $\bar{F}$, integrating by parts and a change of variable yields

$$\int_0^\infty \bar{F}(t)\mathrm{d}t = \int_0^1 \left(\bar{F}\right)^{-1}(s)\mathrm{d}s \tag{4.6}$$

given that $\lim_{t\to\infty} t\bar{F}(t) = \lim_{t\to 0} t\bar{F}(t) = 0$. Eq. (4.6) then gives

$$\int_0^1 \phi(s)\mathrm{d}s = \int_0^\infty \psi(t)\mathrm{d}t = \mathrm{E}(Z) < \infty,$$

and hence $\hat{\mu} \to \mathrm{E}(Z)$ as $n \to \infty$, as claimed.

To show part (ii), write

$$\mathrm{E}(\log Z) = \mathrm{E}\{\log(Z \vee 1)\} + \mathrm{E}\{\log(Z \wedge 1)\} = \mathrm{E}\{\log(Z \vee 1)\} - \mathrm{E}\{\log(1/Z \vee 1)\}.$$

When Condition 3.2 holds, Z is in the domain of attraction of either the Fréchet, the Gumbel or the Weibull distributions. In either case, $\mathrm{E}\{\log(Z \vee 1)\} < \infty$; see Corollary 3.3.32 and Examples 3.3.33 and 3.3.34 of Embrechts et al. (1997). Furthermore, given that $1 - \psi(1/x)$ is the survival function of $1/Z$, Condition 3.1 implies that $1/Z \in \mathcal{M}(\Phi_{1/m})$ and hence $\mathrm{E}\{\log(1/Z \vee 1)\} < \infty$ again using Example 3.3.33 of Embrechts et al. (1997). As in part (i), $\hat{\nu} \to \mathrm{E}(\log Z)$ as $n \to \infty$ then follows directly from

$$\mathrm{E}(\log Z) = \int_0^\infty \psi\{\exp(t)\}\mathrm{d}t = \int_0^1 \log\{\phi(s)\}\mathrm{d}s < \infty,$$

which holds by Eq. (4.6) given that $\psi(e^t)$ is the survival function of $\log Z$. $\square$

Theorems 4.1 and 4.2 below respectively specify the limiting behavior of the processes $\mathbb{A}_n^{\mathrm{CFG}}$ and $\mathbb{A}_n^{\mathrm{P}}$ defined in (4.5). These convergence results require an alpha-mixing (see Definition 2.21) sequence of random variables with a time-invariant Archimax copula. This allows to forgo independence for a form of asymptotic independence in time.

Beforehand, note that the interior of the unit simplex is

$$\mathring{\Delta}_d = \{\boldsymbol{w} \in [0,1]^d : w_1 + \cdots + w_d = 1, w_{(1)} > 0\},$$

where $w_{(1)} = \min(w_1, \ldots, w_d)$. To simplify notation, write, for any $\boldsymbol{x} \in \mathbb{R}_+^d$, $\psi(\boldsymbol{x}) = (\psi(x_1), \ldots, \psi(x_d))$. Furthermore, for any compact subset $\mathcal{K}$ of $\mathring{\Delta}_d$, let $\mathcal{C}(\mathcal{K})$ denote the space of continuous functions on $\mathcal{K}$ equipped with the supremum norm. For a d-variate copula C, let α be a C-Brownian bridge as defined in Chapter 2, (see Equation (2.17)) and recall the definition of the corresponding process $\mathbb{C}$ from Equation (2.18).

Theorem 4.1. *Suppose that $\boldsymbol{X}_1, \boldsymbol{X}_2, \ldots$ is a stationary, alpha-mixing sequence with $\alpha^{[\boldsymbol{X}]}(k) = O(a^k)$, as $k \to \infty$, for some $a \in (0,1)$. Suppose that the marginals of the stationary distribution are continuous and the corresponding copula $C = C_{\psi,\ell} = C_{\psi,A}$ is Archimax and follows the assumptions of Proposition 3.2. Then for any compact set $\mathcal{K} \subset \mathring{\Delta}_d$, $\mathbb{A}_n^{\mathrm{CFG}} \rightsquigarrow \mathbb{A}^{\mathrm{CFG}}$ as $n \to \infty$ in $\mathcal{C}(\mathcal{K})$, where for any $\boldsymbol{w} \in \mathring{\Delta}_d$,*

$$\mathbb{A}^{\mathrm{CFG}}(\boldsymbol{w}) = A(\boldsymbol{w}) \int_0^1 \mathbb{C}[\psi\{-\boldsymbol{w}\log(u)\}] \frac{du}{u \log u}.$$

Theorem 4.2. *Under the assumptions of Theorem 4.1 and the requirement that $s > 2$ when Condition 3.2 (a) holds, one has that, for any compact set $\mathcal{K} \subset \mathring{\Delta}_d$, $\mathbb{A}_n^{\mathrm{P}} \rightsquigarrow \mathbb{A}^{\mathrm{P}}$ as $n \to \infty$ in $\mathcal{C}(\mathcal{K})$, where for any $\boldsymbol{w} \in \mathring{\Delta}_d$,*

$$\mathbb{A}^{\mathrm{P}}(\boldsymbol{w}) = \frac{-A^2(\boldsymbol{w})}{\mathrm{E}(Z)} \int_0^1 \mathbb{C}[\psi\{-\boldsymbol{w}\log(u)\}] \frac{du}{u}.$$

First observe that the conditions of Theorem 4.2 are stronger than those of Theorem 4.1; this was further investigated in Chapter 3. Also note that the generator given, for all $x \geq 0$, by $\psi(x) = e^{-x}$ is completely monotone and satisfies Conditions 3.1 and 3.2 (b) and is such that $-\log(\psi)$ is linear. Hence, Theorems 4.1 and 4.2 remain valid in the special case when C is an extreme-value copula. Finally, note that because of Lemma 4.1, the asymptotic behavior of the endpoint corrected versions of the CFG and Pickands-type estimators is the same, as stated below.

Corollary 4.1. *Theorems 4.1 and 4.2 also hold when $\mathbb{A}_n^{CFG}$ and $\mathbb{A}_n^{\mathrm{P}}$ are respectively replaced by $\mathbb{A}_{n,c}^{\mathrm{CFG}} = \sqrt{n}\,(A_{n,c}^{\mathrm{CFG}} - A)$ and $\mathbb{A}_{n,c}^{\mathrm{P}} = \sqrt{n}\,(A_{n,c}^{\mathrm{P}} - A)$.*

4.2 Proofs of Theorems 4.1 and 4.2

In this section, Theorems 4.1 and 4.2 are proved. To ease the reading, the main arguments are presented in Section 4.2.1. As will be seen therein, the proofs hinge on Proposition 4.2. Auxiliary results are then gathered in Section 4.2.2, with Proposition 4.2 being subsequently proved in two parts in Sections 4.2.3 and 4.2.4.

4.2.1 Outline of the main arguments

To establish weak convergence of $\mathbb{A}_n^{\mathrm{CFG}}$ and $\mathbb{A}_n^{\mathrm{P}}$, the weak convergence of the empirical copula process with respect to weighted metrics established by Berghaus et al. (2017) is used. The result, Theorem 2.2 in said paper, is also reported in Chapter 2 as Theorem 2.12.

Following Genest and Segers (2009), we introduce the processes defined, for any $\boldsymbol{w} \in \Delta_d$, by

$$\mathbb{B}_n^{\mathrm{CFG}}(\boldsymbol{w}) = \sqrt{n}\left\{\log A_n^{\mathrm{CFG}}(\boldsymbol{w}) - \log A(\boldsymbol{w})\right\},$$
$$\mathbb{B}_n^{\mathrm{P}}(\boldsymbol{w}) = \sqrt{n}\left\{1/A_n^{\mathrm{P}}(\boldsymbol{w}) - 1/A(\boldsymbol{w})\right\}.$$

The next lemma establishes that these processes are functionals of the empirical copula process previously defined in (2.16) by $\hat{\mathbb{C}}_n(\boldsymbol{u}) = \sqrt{n}\,\{\hat{C}_n(\boldsymbol{u}) - C(\boldsymbol{u})\}$ for any $\boldsymbol{u} \in [0,1]^d$, where $\hat{C}_n(\boldsymbol{u}) = n^{-1}\sum_{i=1}^n \prod_{j=1}^d \mathbf{1}(\hat{U}_{ij} \le u_j)$ denotes the rank-based empirical copula defined in (2.14) via the pseudo-observations $\hat{U}_{ij}$ as specified in (2.13).

Lemma 4.2. *Fix an arbitrary $\boldsymbol{w} \in \Delta_d$. Then, provided $\mathrm{E}(\log Z)$ exists,*

$$\mathbb{B}_n^{\mathrm{CFG}}(\boldsymbol{w}) = \int_0^1 \hat{\mathbb{C}}_n[\psi\{-\boldsymbol{w}\log(u)\}]\,\frac{du}{u\log u}.$$

Furthermore, provided $\mathrm{E}(Z)$ exists,

$$\mathbb{B}_n^{\mathrm{P}}(\boldsymbol{w}) = \frac{1}{\mathrm{E}(Z)}\int_0^1 \hat{\mathbb{C}}_n[\psi\{-\boldsymbol{w}\log(u)\}]\,\frac{du}{u}.$$

Proof. Using the fact that $\log(t) = \int_0^\infty \{\mathbb{1}(x \le t) - \mathbb{1}(x \le 1)\}\, x^{-1} dx$, for $\boldsymbol{w} \in \Delta_d$, write

$$\begin{aligned}
\mathbb{B}_n^{\mathrm{CFG}}(\boldsymbol{w}) &= -\sqrt{n}\left\{-E\log Z + \frac{1}{n}\sum_{i=1}^n \log\hat{\xi}_i(\boldsymbol{w}) + E\log Z - E\log\xi(\boldsymbol{w})\right\} \\
&= -\sqrt{n}\Bigg(\frac{1}{n}\sum_{i=1}^n \int_0^\infty \left[\mathbb{1}\{x \le \hat{\xi}_i(\boldsymbol{w})\} - \mathbb{1}\{x \le 1\}\right]\frac{dx}{x} \\
&\qquad\qquad - E\int_0^\infty [\mathbb{1}\{x \le \xi(\boldsymbol{w})\} - \mathbb{1}\{x \le 1\}]\,\frac{dx}{x}\Bigg) \\
&= -\sqrt{n}\Bigg(\int_0^\infty \left[\frac{1}{n}\sum_{i=1}^n \mathbb{1}\{x \le \hat{\xi}_i(\boldsymbol{w})\} - \mathbb{1}\{x \le 1\}\right]\frac{dx}{x} \\
&\qquad\qquad - \int_0^\infty [\mathbb{P}\{x \le \xi(\boldsymbol{w})\} - \mathbb{1}\{x \le 1\}]\,\frac{dx}{x}\Bigg) \\
&= -\sqrt{n}\int_0^\infty \Bigg[\frac{1}{n}\sum_{i=1}^n \mathbb{1}\{\hat{U}_{i1} \le \psi(w_1 x), \ldots, \hat{U}_{id} \le \psi(w_d x)\} \\
&\qquad\qquad - \mathbb{P}\left\{U_{i1} \le \psi(w_1 x), \ldots, U_{id} \le \psi(w_d x)\right\}\Bigg]\frac{dx}{x} \\
&= -\int_0^\infty \sqrt{n}\left[\hat{C}_n\{\psi(\boldsymbol{w}x)\} - C\{\psi(\boldsymbol{w}x)\}\right]\frac{dx}{x} \\
&= \int_0^1 \hat{\mathbb{C}}_n[\psi\{-\boldsymbol{w}\log(u)\}]\frac{du}{u\log u}\ .
\end{aligned}$$

Similarly, for the Pickands-type estimator, for $\boldsymbol{w} \in \Delta_d$,

$$\begin{aligned}
\mathbb{B}_n^{\mathrm{P}}(\boldsymbol{w}) &= \sqrt{n}\left\{\frac{\sum_{i=1}^n \hat{\xi}_i(\boldsymbol{w})}{nE(Z)} - \frac{A(\boldsymbol{w})}{E(Z)}\right\} \\
&= \frac{\sqrt{n}}{E(Z)}\int_0^\infty \left[\frac{1}{n}\sum_{i=1}^n \mathbb{1}\{\hat{\xi}_i(\boldsymbol{w}) \ge x\}dx - E\{\xi(\boldsymbol{w})\}\right] dx \\
&= \frac{\sqrt{n}}{E(Z)}\int_0^\infty \left[\frac{1}{n}\sum_{i=1}^n \mathbb{1}\{\hat{\xi}_i(\boldsymbol{w}) \ge x\} - \mathbb{P}\{\xi(\boldsymbol{w}) > x\}\right] dx \\
&= \frac{1}{E(Z)}\int_0^\infty \sqrt{n}\left[\hat{C}_n\{\psi(\boldsymbol{w}x)\} - C\{\psi(\boldsymbol{w}x)\}\right] dx \\
&= \frac{1}{E(Z)}\int_0^1 \hat{\mathbb{C}}_n[\psi\{-\boldsymbol{w}\log(u)\}]\frac{du}{u}\ .
\end{aligned}$$

□

Recall that the required existence of the expectations $\mathrm{E}(\log Z)$ and $\mathrm{E}(Z)$ is treated in Lemma 4.1 and is satisfied under the assumptions of Theorems 4.1 and 4.2, respectively. Weak convergence of $\mathbb{B}_n^{\mathrm{CFG}}$ and $\mathbb{B}_n^{\mathrm{P}}$ is established next. The proof is provided in Sections 4.2.3 and 4.2.4.

Proposition 4.2. *Let $\mathcal{K}$ be any compact subset of $\mathring{\Delta}_d$.*

(a) *Under the assumptions of Theorem 4.1,* $\mathbb{B}_n^{\mathrm{CFG}} \rightsquigarrow \mathbb{B}^{\mathrm{CFG}}$ *as* $n \to \infty$ *in* $\mathcal{C}(\mathcal{K})$*, where for any* $\boldsymbol{w} \in \mathring{\Delta}_d$*,*

$$\mathbb{B}^{\mathrm{CFG}}(\boldsymbol{w}) = \int_0^1 \mathbb{C}[\psi\{-\boldsymbol{w}\log(u)\}]\,\frac{du}{u\log u}\,.$$

(b) *Under the assumptions of Theorem 4.2,* $\mathbb{B}_n^{\mathrm{P}} \rightsquigarrow \mathbb{B}^{\mathrm{P}}$ *as* $n \to \infty$ *in* $\mathcal{C}(\mathcal{K})$*, where for any* $\boldsymbol{w} \in \mathring{\Delta}_d$*,*

$$\mathbb{B}^{\mathrm{P}}(\boldsymbol{w}) = \frac{1}{\mathrm{E}(Z)}\int_0^1 \mathbb{C}[\psi\{-\boldsymbol{w}\log(u)\}]\,\frac{du}{u}.$$

The validity of Theorem 4.1 now follows directly from Proposition 4.2 (a) and Theorem 3.9.4 of van der Vaart and Wellner (1996), given that the map $\eta : \mathcal{C}(\mathcal{K}) \to \mathcal{C}(\mathcal{K})$ defined by $\eta(f) = \exp(f)$ is Hadamard differentiable. Similarly, Theorem 4.2 is a direct consequence of Proposition 4.2 (b) and Slutsky's Lemma, as for any $\boldsymbol{w} \in \Delta_d$,

$$\mathbb{A}_n^{\mathrm{P}}(\boldsymbol{w}) = \frac{-A^2\mathbb{B}_n^{\mathrm{P}}(\boldsymbol{w})}{1 + n^{-1/2}A(\boldsymbol{w})\mathbb{B}_n^{\mathrm{P}}(\boldsymbol{w})}.$$

Remark 4.1. *Theorems 4.1 and 4.2 can in fact be shown to hold for any compact subset* $\mathcal{K}$ *of* $\Delta_d^* = \{\boldsymbol{w} \in [0,1]^d : w_1 + \cdots + w_d = 1, w_{(d)} < 1\}$*, where* $w_{(d)} = \max(w_1, \ldots, w_d)$*. Such sets allow for several components of* $\boldsymbol{w}$ *to be equal to zero. Proposition 4.2 can be proved as follows. Let* $\mathcal{K}$ *be any compact subset of* Δ_d^**. For any* $\boldsymbol{w} = (w_1, \ldots, w_d) \in \mathcal{K}$*, let* $\boldsymbol{w}^\star$ *be the subvector consisting of its non-zero components. Thus* $\boldsymbol{w}^\star$ *is a* $d^\star$*-dimensional vector, with* $d^\star \le d$*, and*

$$\mathbb{B}_n^{\mathrm{CFG}}(\boldsymbol{w}) = -\int_0^\infty \hat{\mathbb{C}}_n^\star\{\psi(\boldsymbol{w}^\star x)\}\frac{dx}{x},\ \mathbb{B}_n^{\mathrm{P}}(\boldsymbol{w}) = \frac{1}{\mathrm{E}(Z)}\int_0^\infty \hat{\mathbb{C}}_n^\star\{\psi(\boldsymbol{w}^\star x)\}dx,$$

where $\hat{\mathbb{C}}_n^\star = \sqrt{n}\,(\hat{C}_n^\star - C^\star)$*. Note that* $C^\star = C_{\psi,\ell^\star}$ *has the same Archimedean generator* ψ *as* C*, and the marginal stdf* $\ell^\star$ *defined as the original* ℓ *with zero arguments corresponding to the zeros of* $\boldsymbol{w}$*. It is then possible to find* $K \in \mathbb{N}$ *such that* $\mathcal{K} \subset B_{1/K} = \{\boldsymbol{w} \in [0,1]^d : w_1 + \cdots + w_d = 1, w_{(1)}^\star \ge 1/K\}$*, where* $w_{(1)}^\star = \min\{w_j : w_j > 0\}$*. The rest of the proof is identical to that of Proposition 4.2. Extending the weak convergence to the entire unit simplex* Δ_d *would require a different approach, and it remains to be seen whether such an extension is possible at all.*

4.2.2 Auxiliary results

In the following, lemmas that are used in the proof of Proposition 4.2 are stated and proved.

Lemma 4.3. *Suppose that* ψ *is a 2-monotone Archimedean generator. Then for any* $K \in \mathbb{N}$ *and* $c \in (0, 1/K)$*, there exists* $N_K \in \mathbb{N}$ *so that for all* $n \ge N_K$*,*

$$\psi\left\{K\phi\left(1 - \frac{c}{n}\right)\right\} > \frac{n}{n+1}.$$

Proof. Let N_K be such that for all $n \geq N_K$, $c < n/\{K(n+1)\}$. Fix an arbitrary $n \geq N_K$ and define, for all $x \geq 0$, $\psi_L(x) = \max(1-x, 0)$ and observe that ψ_L is a 2-monotone Archimedean generator with inverse given, for all $x \in [0,1]$, by $\phi_L(x) = 1-x$. Because ψ is convex, the function $f = \phi_L \circ \psi$ on $[0, \infty)$ is concave and such that $f(0) = 1-\psi(0) = 0$. From Lemma 4.4.3 of Nelsen (2006), f is subadditive. The latter property means that for all $x, y \in [0, \infty)$, $f(x+y) \leq f(x) + f(y)$. Successive application of this inequality yields that for all $x \in [0, \infty)$,

$$f(Kx) \leq Kf(x).$$

Because ψ_L is non-increasing, applying it on both sides gives $\psi_L \circ f(Kx) \geq \psi_L\{Kf(x)\}$. Given that $\psi_L \circ f = \psi$ one has, upon setting $x = \phi(1 - c/n)$,

$$\psi\left\{K\phi\left(1-\frac{c}{n}\right)\right\} \geq \psi_L\left\{K\phi_L\left(1-\frac{c}{n}\right)\right\} = \max\left(1-\frac{Kc}{n}, 0\right) = 1-\frac{Kc}{n},$$

where the last equality follows from the fact that $Kc < 1$ by assumption. Clearly, $1 - (Kc/n) > n/(n+1)$ given that $c < n/\{K(n+1)\}$. $\square$

Lemma 4.4. *(i) If Condition 3.2 holds, then for any $\omega \in (0, 1/2)$ and $a \in (0, x_\psi)$, $\int_a^{x_\psi} \{\psi(x)\}^\omega / x \, dx$ is finite.*

(ii) If Condition 3.2 (a) holds with $s > 2$, then for any $\omega \in (1/s, 1/2)$ and any $a > 0$, $\int_a^\infty \{\psi(x)\}^\omega dx$ is finite.

(iii) If Condition 2 (b) or (c) holds, then for any $\omega \in (0, 1/2)$ and any $a \in (0, x_\psi)$, $\int_a^{x_\psi} \{\psi(x)\}^\omega dx$ is finite.

Proof. (i) If Condition 3.2 (a) holds, $x_\psi = \infty$ and the integrand has index of regular variation $-s\omega - 1 < -1$; the integral is thus finite by Karamata's Theorem (Embrechts et al., 1997, Theorem A3.6). If Condition 3.2 (b) holds and $x_\psi = \infty$, then ψ is rapidly varying and the result follows from Theorem A3.12 (a) of Embrechts et al. (1997). If Condition 3.2 (b) holds and $x_\psi < \infty$ or Condition (c) is satisfied, then $x_\psi < \infty$ and the integrand is bounded on $[a, x_\psi]$.
(ii) Given that the integrand is regularly varying with index $-s\omega < -1$, the result follows from Karamata's Theorem, as in (i).
(iii) In this case, the result follows from Theorem A3.12 (a) of Embrechts et al. (1997) if Condition 3.2 (b) holds and $x_\psi = \infty$, and from fact that $x_\psi < \infty$ otherwise. $\square$

Lemma 4.5. *(i) If Condition 3.2 holds, then for any $c \in (0, 1)$,*

$$\lim_{n\to\infty} \sqrt{n} \int_{\phi(c/n)}^{x_\psi} \frac{\psi(x)}{x} dx = 0 \, .$$

(ii) If either Condition 3.2 (a) with $s > 2$, (b) or (c) holds, then for any $c \in (0, 1)$,

$$\lim_{n\to\infty} \sqrt{n} \int_{\phi(c/n)}^{x_\psi} \psi(x) dx = 0 \, .$$

(iii) If Condition 3.1 holds, then for any $c \in (0,1)$ *and* $\ell \in \{1,2\}$,

$$\lim_{n\to\infty} \sqrt{n} \int_{1/\{K\phi(1-c/n)\}}^{\infty} \frac{1-\psi(1/x)}{x^\ell} dx = 0.$$

Proof. (i) If Condition 3.2 (a) holds, $x_\psi = \infty$. By Karamata's Theorem the integral is a regularly varying function of $\phi(c/n)$ with index $-s$. For some slowly varying function L,

$$\sqrt{n} \int_{\phi(c/n)}^{\infty} \frac{\psi(x)}{x} dx = \sqrt{n} \left\{\phi(c/n)\right\}^{-s} L\left\{\phi(c/n)\right\} .$$

Due to the regular variation of ϕ at zero, there exists a slowly varying function L^* such that

$$\sqrt{n} \left\{\phi(c/n)\right\}^{-s} L\left\{\phi(c/n)\right\} = \sqrt{n} \left\{(n/c)^{1/s} L^*(n/c)\right\}^{-s} L\left\{\phi(c/n)\right\} \tag{4.7}$$

which may be written as $(c/\sqrt{n})L^\dagger(n)$, where $L^\dagger(n) = L^*(n/c)^{-s} L\{\phi(c/n)\}$ is a slowly varying function of n, see, e.g., Proposition 0.8 (iv) of Resnick (1987). Consequently, the left-hand side of (4.7) converges to zero as $n \to \infty$.

If Condition 3.2 (b) holds and $x_\psi = \infty$, Theorem A3.12 (b) of Embrechts et al. (1997) implies that

$$\lim_{n\to\infty} \frac{n}{c} \int_{\phi(c/n)}^{\infty} \frac{\psi(x)}{x} dx = 0,$$

from which the result follows at once. Finally, if Condition 3.2 (b) holds and $x_\psi < \infty$ or if Condition 3.2 (c) is satisfied, $x_\psi < \infty$ and $\psi(x) = 0$ for all $x \geq x_\psi$. Because ψ is decreasing,

$$\sqrt{n} \int_{\phi(c/n)}^{x_\psi} \frac{\psi(x)}{x} dx \leq \sqrt{n} \int_{\phi(c/n)}^{x_\psi} \frac{\psi\left\{\phi(c/n)\right\}}{x} dx = \frac{\log x_\psi - \log\{\phi(c/n)\}}{\sqrt{n}/c} .$$

Clearly, the last expression converges to zero as $n \to \infty$.
(ii) If Condition 3.2 (a) holds with $s > 2$, $x_\psi = \infty$ and one can argue as in the proof of (i) using Karamata's Theorem that

$$\sqrt{n} \int_{b_n}^{\infty} \psi(x) dx = n^{1/2+1/s-1} L^{\dagger\dagger}(n),$$

where $L^{\dagger\dagger}$ is slowly varying. Since $1/2 + 1/s - 1 < 0$, the right-hand side converges to 0 as $n \to \infty$. If Condition 3.2 (b) holds and $x_\psi = \infty$, Theorem A3.12 (b) of Embrechts et al. (1997) and the fact that $\phi(1/x)$ is slowly varying (Bingham et al., 1989, Theorem 2.4.7) imply that

$$\lim_{n\to\infty} \sqrt{n} \int_{\phi(c/n)}^{\infty} \psi(x) dx = \lim_{n\to\infty} \frac{c\phi(c/n)}{\sqrt{n}} \frac{\int_{\phi(c/n)}^{\infty} \psi(t) dt}{(c/n)\phi(c/n)} = 0 .$$

If Condition 3.2 (b) holds and $x_\psi < \infty$ or Condition 3.2 (c) is satisfied, then $x_\psi < \infty$. Consequently,

$$\sqrt{n} \int_{\phi(c/n)}^{x_\psi} \psi(x) dx \leq \sqrt{n}(c/n)\{x_\psi - \phi(c/n)\} ;$$

the last expression clearly converges to zero as $n \to \infty$.

(iii) Because for sufficiently large n,

$$0 \le \sqrt{n} \int_{1/\{K\phi(1-c/n)\}}^{\infty} \frac{1-\psi(1/x)}{x^2} dx \le \sqrt{n} \int_{1/\{K\phi(1-c/n)\}}^{\infty} \frac{1-\psi(1/x)}{x} dx,$$

it suffices to consider the case $\ell = 1$. Karamata's Theorem implies that there exists a slowly varying function L_1 such that

$$\sqrt{n} \int_{1/\{K\phi(1-c/n)\}}^{\infty} \frac{1-\psi(1/x)}{x} dx = \sqrt{n} \left\{ K\phi\left(1 - \frac{c}{n}\right)\right\}^{\frac{1}{m}} L_1 \left[\left\{ K\phi\left(1 - \frac{c}{n}\right)\right\}^{-1} \right].$$

Because $\phi(1 - 1/x)$ is regularly varying with index $-m$, there exists a slowly varying function L_2 such that

$$\sqrt{n} \left\{ K\phi\left(1 - \frac{c}{n}\right)\right\}^{\frac{1}{m}} L_1 \left[\left\{ K\phi\left(1 - \frac{c}{n}\right)\right\}^{-1} \right] = \sqrt{n} \left\{ K(n/c)^{-m} L_2(n/c)\right\}^{\frac{1}{m}} L_1 \left[\left\{ K\phi\left(1 - \frac{c}{n}\right)\right\}^{-1} \right] = n^{-1/2} L_3(n),$$

where $L_3(x) = cK^{1/m} L_2(x/c)^{1/m} L_1[\{K\phi(1 - c/x)\}^{-1}]$. As L_3 is slowly varying (Resnick, 1987, Proposition 0.8 (iv)), $n^{-1/2} L_3(n) \to 0$ as $n \to \infty$. □

Remark 4.2. *It emerges from the proofs of Lemma 4.4 and 4.5 that these results remain valid if instead of Condition 3.2 (b) or (c), ψ satisfies the weaker condition that either $x_\psi < \infty$, or that $x_\psi = \infty$ and ψ is rapidly varying as defined, e.g., on p. 83 in Bingham et al. (1989).*

4.2.3 Proof of Proposition 4.2 (a)

Let $\mathcal{K}$ be a compact subset of $\mathring{\Delta}_d$. For an arbitrary $\boldsymbol{w} \in \Delta_d$, set $w_{(1)} = \min_{i=1,\ldots,d} w_i$ and $w_{(d)} = \max_{i=1,\ldots,d} w_i$. Define, for any $k \in \mathbb{N}$, the set $B_{1/k} = \{\boldsymbol{w} \in \Delta_d : w_{(1)} \ge 1/k\}$. Since $\mathcal{K}$ is compact, there exists an integer $K > 1$ such that $\mathcal{K} \subset B_{1/K} \subset \mathring{\Delta}_d$. Next, pick an arbitrary $c \in (0, 1/K^{1/m})$ with m from Condition 3.1, and define

$$a_n = \phi\left(1 - \frac{c}{n}\right), \quad b_n = \phi\left(\frac{c}{n}\right) . \tag{4.8}$$

By Lemma 4.3 and because $c < 1$, there exists $N_K \in \mathbb{N}$ so that for any $n \ge N_K$,

$$c < \frac{n}{n+1} \quad \text{and} \quad \psi\left\{K\phi\left(1 - \frac{c}{n}\right)\right\} > \frac{n}{n+1} . \tag{4.9}$$

Next, for any $i \ge 1$ and $j \in \{1, \ldots, d\}$, let $U_{ij} = F_j(X_{ij})$ and set $\boldsymbol{U}_i = (U_{i1}, \ldots, U_{id})$. Recall from Chapter 2 that the empirical copula and empirical copula process pertaining to the unobservable sequence $\boldsymbol{U}_1, \ldots, \boldsymbol{U}_n$ are given for any $\boldsymbol{u} \in [0, 1]^d$, by $G_n(\boldsymbol{u}) =$

$\frac{1}{n}\sum_{i=1}^n \prod_{j=1}^d \mathbf{1}(U_{ij} \le u_j)$ and $\alpha_n(\boldsymbol{u}) = \sqrt{n}\{G_n(\boldsymbol{u}) - C(\boldsymbol{u})\}$ respectively (see Equation (2.19)). Recall also from Chapter 2, Theorem 2.12, the process defined at any $\boldsymbol{u} \in [0,1]^d$ by $\bar{\mathbb{C}}_n(\boldsymbol{u}) = \alpha_n(\boldsymbol{u}) - \sum_{j=1}^d \dot{C}_j(\boldsymbol{u})\alpha_n(\boldsymbol{u}^{(j)})$.

Before proceeding, recall that for any $\boldsymbol{x} \in \mathbb{R}_+^d$, $\psi(\boldsymbol{x}) = (\psi(x_1), \ldots, \psi(x_d))$ and note the following lemma.

Lemma 4.6. *As $n \to \infty$, $\sup_{\boldsymbol{w} \in B_{1/K}} \int_0^{x_\psi/w_{(d)}} |\bar{\mathbb{C}}_n\{\psi(\boldsymbol{w}x)\} - \hat{\mathbb{C}}_n\{\psi(\boldsymbol{w}x)\}| \frac{dx}{x}$ converges in probability to* 0.

Proof. Using triangle inequality and a_n, b_n as in (4.8), write, for any $\boldsymbol{w} \in B_{1/K}$,

$$\int_0^{x_\psi/w_{(d)}} |\bar{\mathbb{C}}_n\{\psi(\boldsymbol{w}x)\} - \hat{\mathbb{C}}_n\{\psi(\boldsymbol{w}x)\}| \frac{dx}{x} \le \sum_{j=1}^5 I_j(\boldsymbol{w}),$$

where

$$I_1(\boldsymbol{w}) = \int_{a_n/w_{(1)}}^{b_n/w_{(d)}} \left|\hat{\mathbb{C}}_n\{\psi(\boldsymbol{w}x)\} - \bar{\mathbb{C}}_n\{\psi(\boldsymbol{w}x)\}\right| \frac{dx}{x},$$

$$I_2(\boldsymbol{w}) = \int_0^{a_n/w_{(1)}} \left|\hat{\mathbb{C}}_n\{\psi(\boldsymbol{w}x)\}\right| \frac{dx}{x},$$

$$I_3(\boldsymbol{w}) = \int_{b_n/w_{(d)}}^{x_\psi/w_{(d)}} \left|\hat{\mathbb{C}}_n\{\psi(\boldsymbol{w}x)\}\right| \frac{dx}{x}, \quad I_4(\boldsymbol{w}) = \int_0^{a_n/w_{(1)}} \left|\bar{\mathbb{C}}_n\{\psi(\boldsymbol{w}x)\}\right| \frac{dx}{x},$$

$$I_5(\boldsymbol{w}) = \int_{b_n/w_{(d)}}^{x_\psi/w_{(d)}} \left|\bar{\mathbb{C}}_n\{\psi(\boldsymbol{w}x)\}\right| \frac{dx}{x}.$$

In the sequel, we show that for any $p \in \{1, \ldots, 5\}$, $\sup_{\boldsymbol{w} \in B_{1/K}} I_p(\boldsymbol{w}) \to 0$ in probability as $n \to \infty$.

Treatment of I_1. Fix an arbitrary $\boldsymbol{w} \in B_{1/K}$ and introduce, for any $\omega \in (0, 1/2)$, the weight function g_ω from Theorem 2.2 in Berghaus et al. (2017) reported in (2.20). The latter is given at any $\boldsymbol{u} \in [0,1]^d$ by

$$g_\omega(\boldsymbol{u}) = \min\left[\bigwedge_{i=1}^d u_j, \bigwedge_{i=1}^d \left\{1 - \min_{j=1,\ldots,d}(u_1, \ldots, u_{j-1}, u_{j+1}, \ldots, u_d)\right\}\right]^\omega. \tag{4.10}$$

Because $a_n/w_{(1)} < x < b_n/w_{(d)}$ implies that, for all $j \in \{1, \ldots, d\}$, $c/n < \psi(w_j x) < 1 - c/n$, one has

$$\begin{aligned} I_1(\boldsymbol{w}) &= \int_{a_n/w_{(1)}}^{b_n/w_{(d)}} \left|\frac{\hat{\mathbb{C}}_n\{\psi(\boldsymbol{w}x)\}}{g_\omega\{\psi(\boldsymbol{w}x)\}} - \frac{\bar{\mathbb{C}}_n\{\psi(\boldsymbol{w}x)\}}{g_\omega\{\psi(\boldsymbol{w}x)\}}\right| \frac{g_\omega\{\psi(\boldsymbol{w}x)\}}{x} dx \\ &\le S_n \int_0^{x_\psi/w_{(d)}} \frac{g_\omega\{\psi(\boldsymbol{w}x)\}}{x} dx, \end{aligned}$$

where

$$S_n = \sup_{\boldsymbol{u} \in [c/n, 1-c/n]^d} \left|\frac{\hat{\mathbb{C}}_n(\boldsymbol{u})}{g_\omega(\boldsymbol{u})} - \frac{\bar{\mathbb{C}}_n(\boldsymbol{u})}{g_\omega(\boldsymbol{u})}\right|. \tag{4.11}$$

By the first part of Theorem 2.12, S_n converges to 0 in probability as $n \to \infty$. The conditions of the latter Theorem are indeed fulfilled because of Proposition 3.2. To conclude that $\sup_{\boldsymbol{w}\in B_{1/K}} I_1(\boldsymbol{w}) \to 0$ in probability as $n \to \infty$, it thus suffices to show that $\int_0^{x_\psi/w_{(d)}} \frac{g_\omega\{\psi(\boldsymbol{w}x)\}}{x} dx$ is finite. To this end, note that because ψ is decreasing,

$$g_\omega\{\psi(\boldsymbol{w}x)\} \le [\min\{\psi(xw_1), \dots, \psi(xw_d)\}]^\omega = \{\psi(w_{(d)}x)\}^\omega \tag{4.12}$$

and that, since $w_j \le 1$ for all $j \in \{1, \dots, d\}$,

$$g_\omega\{\psi(\boldsymbol{w}x)\} \le [1 - \min\{\psi(xw_1), \dots, \psi(xw_d)\}]^\omega = \{1 - \psi(w_{(d)}x)\}^\omega . \tag{4.13}$$

Choosing an arbitrary $a \in (0, x_\psi)$, one then has

$$\begin{aligned}\int_0^{x_\psi/w_{(d)}} \frac{g_\omega\{\psi(\boldsymbol{w}x)\}}{x} dx \le& \int_0^{a/w_{(d)}} \frac{\{1 - \psi(w_{(d)}x)\}^\omega}{x} dx \\ &+ \int_{a/w_{(d)}}^{x_\psi/w_{(d)}} \frac{\{\psi(w_{(d)}x)\}^\omega}{x} dx = I_{11} + I_{12} < \infty,\end{aligned} \tag{4.14}$$

where

$$I_{11} = \int_{1/a}^\infty \frac{\{1 - \psi(1/x)\}^\omega}{x} dx, \quad I_{12} = \int_a^{x_\psi} \frac{\{\psi(x)\}^\omega}{x} dx. \tag{4.15}$$

Indeed, under Condition 3.1, I_{11} is finite by Karamata's Theorem, since the integrand has index of regular variation $-m\omega - 1$ which is strictly less than -1. Finally, I_{12} is finite under Condition 3.2 by Lemma 4.4 (i).

Treatment of I_2. Without loss of generality, suppose that $n \ge N_K$ so that (4.9) holds. Fix an arbitrary $\boldsymbol{w} \in B_{1/K}$ and observe that from the definition of $B_{1/K}$ one has, for any $x \in (0, a_n/w_{(1)})$ and $j \in \{1, \dots, d\}$,

$$w_j x \le \frac{w_j}{w_{(1)}} \phi\left(1 - \frac{c}{n}\right) \le K\phi\left(1 - \frac{c}{n}\right) .$$

This and (4.9) imply that

$$\psi(w_j x) \ge \psi\{K\phi(1 - c/n)\} > \frac{n}{n+1} .$$

Consequently, for any $x \in (0, a_n/w_{(1)})$, $\hat{C}_n\{\psi(\boldsymbol{w}x)\} = 1$. Using (2.8), one thus has

$$I_2(\boldsymbol{w}) = \sqrt{n} \int_0^{a_n/w_{(1)}} [1 - C\{\psi(\boldsymbol{w}x)\}] \frac{dx}{x} = \sqrt{n} \int_0^{a_n/w_{(1)}} \frac{1 - \psi\{\ell(\boldsymbol{w}x)\}}{x} dx.$$

Because for any $x > 0$, $\ell(\boldsymbol{w}x) = x\ell(\boldsymbol{w})$, $\ell(\boldsymbol{w}) \le 1$, and $w_{(1)} \ge 1/K$ one further has that

$$I_2(\boldsymbol{w}) \le \sqrt{n} \int_{w_{(1)}/a_n}^\infty \frac{1 - \psi(1/x)}{x} dx \le \sqrt{n} \int_{1/(Ka_n)}^\infty \frac{1 - \psi(1/x)}{x} dx .$$

The last term in the above inequality is independent of $\boldsymbol{w}$ and converges to 0 as $n \to \infty$ by Lemma 4.5 (iii).

Treatment of I_3. Without loss of generality, suppose that $n \geq N_K$ so that (4.9) holds. Fix an arbitrary $\boldsymbol{w} \in B_{1/K}$ and observe that if $x \geq b_n/w_{(d)}$, $\psi(xw_{(d)}) \leq c/n < 1/(n+1)$ and consequently $\hat{C}_n\{\psi(\boldsymbol{w}x)\} = 0$. Thus

$$I_3(\boldsymbol{w}) = \sqrt{n} \int_{b_n/w_{(d)}}^{x_\psi/w_{(d)}} C\{\psi(\boldsymbol{w}x)\} \frac{dx}{x} \leq \sqrt{n} \int_{b_n/w_{(d)}}^{x_\psi/w_{(d)}} \frac{\psi(w_{(d)}x)}{x} dx = \sqrt{n} \int_{b_n}^{x_\psi} \frac{\psi(x)}{x} dx \, .$$

The last term in the above inequality is independent of $\boldsymbol{w}$ and converges to 0 as $n \to \infty$ by Lemma 4.5 (i).

Treatment of I_4. Recall the second weight function $\tilde{g}_\omega$ from Berghaus et al. (2017) reproduced in (2.21). Fix an arbitrary $\boldsymbol{w} \in B_{1/K}$, let

$$Z_n = \sup_{\boldsymbol{u} \in [0,1]^d} \left| \frac{\bar{\mathbb{C}}_n(\boldsymbol{u})}{\tilde{g}_\omega(\boldsymbol{u})} \right| \tag{4.16}$$

and observe that

$$\begin{aligned} I_4(\boldsymbol{w}) = \int_0^{a_n/w_{(1)}} \left| \frac{\bar{\mathbb{C}}_n\{\psi(\boldsymbol{w}x)\}}{\tilde{g}_\omega\{\psi(\boldsymbol{w}x)\}} \right| \frac{\tilde{g}_\omega\{\psi(\boldsymbol{w}x)\}}{x} dx &\leq \int_0^{a_n/w_{(1)}} Z_n \frac{\tilde{g}_\omega\{\psi(\boldsymbol{w}x)\}}{x} dx \\ &\leq Z_n \int_0^{Ka_n} \frac{\tilde{g}_\omega\{\psi(\boldsymbol{w}x)\}}{x} dx. \end{aligned}$$

Given that $Z_n \rightsquigarrow \sup_{\boldsymbol{u} \in [0,1]^d} |\mathbb{C}(\boldsymbol{u})/\tilde{g}_\omega(\boldsymbol{u})|$ as $n \to \infty$ by Theorem 2.12, it suffices to prove that

$$\int_0^{Ka_n} \frac{\tilde{g}_\omega\{\psi(\boldsymbol{w}x)\}}{x} dx$$

converges uniformly to 0 as $n \to \infty$. To this end, note that $g_\omega(\boldsymbol{u}) = 0$ occurs either when at least one component of $\boldsymbol{u}$ is equal to 0 or at least $d-1$ components are equal to 1. Given that $a_n \to 0$ as $n \to \infty$, one thus has, for sufficiently large n,

$$\int_0^{Ka_n} \frac{\tilde{g}_\omega\{\psi(\boldsymbol{w}x)\}}{x} dx = \int_0^{Ka_n} \frac{g_\omega\{\psi(\boldsymbol{w}x)\}}{x} dx.$$

Using (4.13), the integral on the right-hand side can be bounded above by

$$\begin{aligned} \int_0^{Ka_n/w_{(d)}} \frac{g_\omega\{\psi(\boldsymbol{w}x)\}}{x} dx &\leq \int_0^{Ka_n/w_{(d)}} \frac{\{1 - \psi(w_{(d)}x)\}^\omega}{x} dx \\ &= \int_{1/(Ka_n)}^{\infty} \frac{\{1 - \psi\left(1/x\right)\}^\omega}{x} dx \, . \end{aligned}$$

The last expression converges to 0 as $n \to \infty$, given that it is bounded above by I_{11} in (4.14), which is finite, and given that $a_n \to 0$ as $n \to \infty$.

Treatment of I_5. Let $\tilde{g}_\omega$ be as in the preceding paragraph concerning I_4. Fix an arbitrary $\boldsymbol{w} \in B_{1/K}$ and note that, using (4.12) and performing a change of variable,

$$\begin{aligned} I_5(\boldsymbol{w}) \ \leq \ Z_n \int_{b_n/w_{(d)}}^{x_\psi/w_{(d)}} \frac{\tilde{g}_\omega\{\psi(\boldsymbol{w}x)\}}{x} dx &= Z_n \int_{b_n/w_{(d)}}^{x_\psi/w_{(d)}} \frac{g_\omega\{\psi(\boldsymbol{w}x)\}}{x} dx \\ &\leq Z_n \int_{b_n}^{x_\psi} \frac{\{\psi(x)\}^\omega}{x} dx. \end{aligned}$$

The claim follows since $\int_{b_n}^{x_\psi} \frac{\{\psi(x)\}^\omega}{x} dx \to 0$ as $n \to \infty$ by Lemma 4.4 (i) given that $b_n \to x_\psi$ as $n \to \infty$. $\square$

Returning to the proof of Proposition 4.2 (a), fix an arbitrary $\boldsymbol{w} \in B_{1/K}$ and observe that from Lemma 4.2 and the fact that $\hat{C}_n\{\psi(\boldsymbol{w}x)\} = C\{\psi(\boldsymbol{w}x)\} = 0$ whenever $x > x_\psi/w_{(d)}$,

$$\mathbb{B}_n^{\mathrm{CFG}}(\boldsymbol{w}) = -\int_0^{x_\psi/w_{(d)}} \hat{\mathbb{C}}_n\{\psi(\boldsymbol{w}x)\}\frac{dx}{x}.$$

Now introduce the process $\bar{\mathbb{B}}_n^{\mathrm{CFG}}$ given, for any $\boldsymbol{w} \in \Delta_d$, by

$$\bar{\mathbb{B}}_n^{\mathrm{CFG}}(\boldsymbol{w}) = -\int_0^{x_\psi/w_{(d)}} \bar{\mathbb{C}}_n\{\psi(\boldsymbol{w}x)\}\frac{dx}{x}.$$

From Lemma 4.6, it follows that $\sup_{\boldsymbol{w}\in B_{1/K}} |\mathbb{B}_n^{\mathrm{CFG}}(\boldsymbol{w}) - \bar{\mathbb{B}}_n^{\mathrm{CFG}}(\boldsymbol{w})|$ converges to zero in probability. It thus remains to show that $\bar{\mathbb{B}}_n^{\mathrm{CFG}} \rightsquigarrow \mathbb{B}^{\mathrm{CFG}}$ in $\mathcal{C}(B_{1/K})$ as $n \to \infty$. To do so, consider the map

$$\begin{aligned} \Gamma \ : \ \big(\ell^\infty([0,1]^d), ||\cdot||_{\tilde{g}_\omega}\big) &\longmapsto \big(\ell^\infty(B_{1/K}), ||\cdot||_\infty\big) \\ f \quad &\longmapsto \left\{ \boldsymbol{w} \ \mapsto -\int_0^{x_\psi/w_{(d)}} f\{\psi(\boldsymbol{w}x)\}\frac{dx}{x} \right\}, \end{aligned} \tag{4.17}$$

where $||f||_{\tilde{g}_\omega} = \sup_{\boldsymbol{u}\in[0,1]^d} |f(\boldsymbol{u})/\tilde{g}_\omega(\boldsymbol{u})|$. Let f_1, f_2 be arbitrary functions in $(\ell^\infty([0,1]^d), ||\cdot||_{\tilde{g}_\omega})$. Then

$$\begin{aligned} \sup_{\boldsymbol{w}\in B_{1/K}} |\Gamma(f_1) - \Gamma(f_2)| &= \sup_{\boldsymbol{w}\in B_{1/K}} \left| -\int_0^{x_\psi/w_{(d)}} \frac{f_1\{\psi(\boldsymbol{w}x)\} - f_2\{\psi(\boldsymbol{w}x)\}}{\tilde{g}_\omega\{\psi(\boldsymbol{w}x)\}} \frac{\tilde{g}_\omega\{\psi(\boldsymbol{w}x)\}}{x} dx \right| \\ &\le \sup_{\boldsymbol{w}\in B_{1/K}} \left| \int_0^{x_\psi/w_{(d)}} ||f_1 - f_2||_{\tilde{g}_\omega} \frac{\tilde{g}_\omega\{\psi(\boldsymbol{w}x)\}}{x} dx \right| \\ &\le ||f_1 - f_2||_{\tilde{g}_\omega}(I_{11} + I_{12}), \end{aligned}$$

where the last inequality follows from (4.14). The map Γ is thus Lipschitz. Theorem 2.12 and the Continuous Mapping Theorem then imply that $\bar{\mathbb{B}}_n^{\mathrm{CFG}} = \Gamma(\bar{\mathbb{C}}_n) \rightsquigarrow \Gamma(\mathbb{C}) = \mathbb{B}^{\mathrm{CFG}}$ as $n \to \infty$ weakly in $\ell^\infty(B_{1/K})$. Since $\mathbb{B}^{\mathrm{CFG}}$ has continuous paths on $B_{1/K}$, the convergence takes place on $\mathcal{C}(B_{1/K})$. $\square$

4.2.4 Proof of Proposition 4.2 (b)

The proof of Proposition 4.2 (b) is similar to the proof of part (a) detailed in Section 4.2.3. For the sake of brevity, only the differences are pointed out.

Let $\mathcal{K}$ be a compact subset of $\mathring{\Delta}_d$. Let $B_{1/K}$ and c be as in Section 4.2.3 and a_n, b_n as in (4.8). Furthermore, assume without loss of generality that n is sufficiently large so that (4.9) holds. Finally, recall the weight function g_ω given in (4.10) for some arbitrary fixed $\omega \in (0, 1/2)$; if Condition 3.2 (a) holds, $\omega \in (0, 1/2)$ must in addition be such that $s\omega > 1$. The following result is the analogue of Lemma 4.6.

Lemma 4.7. *As $n \to \infty$, $\sup_{\boldsymbol{w} \in B_{1/K}} \int_0^{x_\psi/w_{(d)}} |\bar{\mathbb{C}}_n\{\psi(\boldsymbol{w}x)\} - \hat{\mathbb{C}}_n\{\psi(\boldsymbol{w}x)\}|dx$ converges in probability to* 0.

Proof. Fix an arbitrary $\boldsymbol{w} \in B_{1/K}$. Then

$$\int_0^{x_\psi/w_{(d)}} |\bar{\mathbb{C}}_n\{\psi(\boldsymbol{w}x)\} - \hat{\mathbb{C}}_n\{\psi(\boldsymbol{w}x)\}|dx \leq \sum_{j=1}^{5} I_j(\boldsymbol{w}),$$

where

$$I_1(\boldsymbol{w}) = \int_{a_n/w_{(1)}}^{b_n/w_{(d)}} \left|\hat{\mathbb{C}}_n\{\psi(\boldsymbol{w}x)\} - \bar{\mathbb{C}}_n\{\psi(\boldsymbol{w}x)\}\right| dx \ ,$$

$$I_2(\boldsymbol{w}) = \int_0^{a_n/w_{(1)}} \left|\hat{\mathbb{C}}_n\{\psi(\boldsymbol{w}x)\}\right| dx,$$

$$I_3(\boldsymbol{w}) = \int_{b_n/w_{(d)}}^{x_\psi/w_{(d)}} \left|\hat{\mathbb{C}}_n\{\psi(\boldsymbol{w}x)\}\right| dx, \quad I_4(\boldsymbol{w}) = \int_0^{a_n/w_{(1)}} \left|\bar{\mathbb{C}}_n\{\psi(\boldsymbol{w}x)\}\right| dx,$$

$$I_5(\boldsymbol{w}) = \int_{b_n/w_{(d)}}^{x_\psi/w_{(d)}} \left|\bar{\mathbb{C}}_n\{\psi(\boldsymbol{w}x)\}\right| dx.$$

To prove the claim, we show that for any $p \in \{1 \ldots, 5\}$,
$\sup_{\boldsymbol{w} \in B_{1/K}} I_p(\boldsymbol{w}) \to 0$ in probability as $n \to \infty$.

Treatment of I_1. Define S_n as in (4.11) and observe that

$$I_1(\boldsymbol{w}) \leq S_n \int_{a_n/w_{(1)}}^{b_n/w_{(d)}} g_\omega\{\psi(\boldsymbol{w}x)\}dx \leq S_n \int_0^{x_\psi/w_{(d)}} g_\omega\{\psi(\boldsymbol{w}x)\}dx \ .$$

For an arbitrary $a \in (0, x_\psi)$ one further has, using (4.12) and (4.13) and the fact that $w_{(d)} \geq 1/d$,

$$\int_0^{x_\psi/w_{(d)}} g_\omega\{\psi(\boldsymbol{w}x)\}dx \leq d \int_0^a \{1 - \psi(x)\}^\omega dx + d \int_a^{x_\psi} \psi(x)^\omega dx. \tag{4.18}$$

The upper bound in the preceding display is finite; this follows from Lemma 4.4 (ii)–(iii) and the fact that $\{1 - \psi(x)\}^\omega$ is bounded on $[0, a]$. Given that S_n converges to 0 in probability as $n \to \infty$ by Theorem 2.12, $\sup_{\boldsymbol{w} \in B_{1/K}} I_1(\boldsymbol{w}) \to 0$ in probability as $n \to \infty$, as claimed.

Treatment of I_2. Fix an arbitrary $\boldsymbol{w} \in B_{1/K}$. Using the same arguments as in the paragraph concerning the treatment of I_2 in the proof of Lemma 4.6, one has that

$$I_2(\boldsymbol{w}) \leq \sqrt{n} \int_0^{a_n/w_{(1)}} \{1 - \psi(x)\}dx \leq \frac{\sqrt{n}}{w_{(1)}} \phi(1 - c/n) \leq K\sqrt{n}\phi(1 - c/n) \ .$$

Given that $\sqrt{x}\phi(1 - c/x)$ is regularly varying of index $1/2 - m < 0$, the expression on the right-hand side converges to 0 as $n \to \infty$.

Treatment of I_3. Fix an arbitrary $\boldsymbol{w} \in B_{1/K}$. Using the same arguments as in the paragraph concerning the treatment of I_3 in the proof of Lemma 4.6 and the fact that $w_{(d)} \geq 1/d$, one has that $I_3(\boldsymbol{w}) \leq d\sqrt{n} \int_{b_n}^{x_\psi} \psi(x)dx$. The upper bound converges to 0 as $n \to \infty$ by Lemma 4.5 (ii).

Treatment of I_4. Fix an arbitrary $\boldsymbol{w} \in B_{1/K}$. Arguing as in the paragraph concerning the treatment of I_4 in the proof of Lemma 4.6 and using the fact that $w_{(d)} \geq 1/d$ one has that

$$I_4(\boldsymbol{w}) \leq Z_n \int_0^{Ka_n} g_\omega\{\psi(\boldsymbol{w}x)\}dx \leq Z_n d \int_0^{Ka_n} \{1 - \psi(x)\}^\omega dx \; .$$

The upper bound converges in probability to 0 as $n \to \infty$, given that Z_n converges in distribution by Theorem 2.12, and $\int_0^{Ka_n} \{1 - \psi(x)\}^\omega dx \to 0$ as $n \to \infty$, given that $a_n \to 0$ as $n \to \infty$.

Treatment of I_5. Fix an arbitrary $\boldsymbol{w} \in B_{1/K}$. Arguing as in the paragraph concerning the treatment of I_5 in the proof of Lemma 4.6, one has that

$$I_5(\boldsymbol{w}) \leq Z_n d \int_{b_n}^{x_\psi} \{\psi(x)\}^\omega dx \; .$$

As in the preceding paragraph, the claim follows from the fact that
$\int_{b_n}^{x_\psi} \psi(x)dx \to 0$ as $n \to \infty$ given that $b_n \to x_\psi$ as $n \to \infty$. $\square$

Returning to the proof of Proposition 4.2 (b), introduce the process $\bar{\mathbb{B}}_n^{\mathrm{P}}$ given, for all $\boldsymbol{w} \in \Delta_d$, by

$$\bar{\mathbb{B}}_n^{\mathrm{P}}(\boldsymbol{w}) = \frac{1}{E(Z)} \int_0^{x_\psi/w_{(d)}} \bar{\mathbb{C}}_n\{\psi(\boldsymbol{w}x)\}dx \; .$$

From Lemma 4.2 one has that

$$\mathbb{B}_n^{\mathrm{P}}(\boldsymbol{w}) = \frac{1}{E(Z)} \int_0^{x_\psi/x_{(d)}} \hat{\mathbb{C}}_n\{\psi(\boldsymbol{w}x)\}dx \; ,$$

and Lemma 4.7 implies that $\sup_{\boldsymbol{w} \in B_{1/K}} |\mathbb{B}_n^{\mathrm{P}}(\boldsymbol{w}) - \bar{\mathbb{B}}_n^{\mathrm{P}}(\boldsymbol{w})| \to 0$ in probability as $n \to \infty$. As in the proof of Proposition 4.2 (b), one can establish that $\bar{\mathbb{B}}_n^{\mathrm{P}} \rightsquigarrow \mathbb{B}^{\mathrm{P}}$ as $n \to \infty$ in $\mathcal{C}(B_{1/K})$ using Theorem 2.12 and the Continuous Mapping Theorem featuring the map

$$\begin{aligned} \Gamma \; : \; \big(\ell^\infty([0,1]^d), ||\cdot||_{\tilde{g}_\omega}\big) &\longmapsto \big(\ell^\infty(B_{1/K}), ||\cdot||_\infty\big) \\ f \quad &\longmapsto \left\{ \boldsymbol{w} \mapsto \int_0^{x_\psi/w_{(d)}} f\{\psi(\boldsymbol{w}x)\}dx \right\} \; , \end{aligned}$$

which is easily shown to be Lipschitz. $\square$

Table 4.1: Archimedean generators and stdfs used in the simulation study in Section 4.3.

	Archimedean generators			
Family	$\psi_\theta(x)$	$\mathcal{O}$	Cond. 3.1	Cond. 3.2
Clayton	$(1+\theta x)^{-1/\theta}$	$(0,\infty)$	✓ $(m=1)$	✓ (a; $s=1/\theta$)
Frank	$-(1/\theta)\log\{1+e^{-x}(e^{-\theta}-1)\}$	$\mathbb{R}$	✓ $(m=1)$	✓ (b)
Gumbel	$\exp(-x^{1/\theta})$	$[1,\infty)$	✓ $(m=\theta)$	✓ (b)
Joe	$1-\{1-e^{-x}\}^{1/\theta}$	$[1,\infty)$	✓ $(m=\theta)$	✓ (b)
	Stable tail dependence functions			
Family	$\ell(x_1,\dots,x_d)$		Parameters	
LG	$(x_1^\varrho+\dots+x_d^\varrho)^{\frac{1}{\varrho}}$		$\varrho\in[1,\infty)$	
NSD	$\frac{\Gamma(\alpha_1+\cdots+\alpha_d-\rho)}{\Gamma(\alpha_1+\cdots+\alpha_d)}\,\mathrm{E}\Big\{\max\limits_{1\le j\le d}\Big(\frac{x_jD_j^{-\rho}\Gamma(\alpha_j)}{\Gamma(\alpha_j-\rho)}\Big)\Big\}$		$(D_1,\dots,D_d)\sim\text{Dirichlet}(\alpha_1,\dots,\alpha_d)$ $\alpha_1,\dots,\alpha_d>0$, $\rho\in(0,\min(\alpha_1,\dots,\alpha_d))$	
DSM	$d\sum_{\boldsymbol{w}\in\mathcal{W}}\max(x_1w_1,\dots,x_dw_d)$		$\mathcal{W}$ is a finite subset of Δ_d with cardinality m given in (A.1)–(A.3) in Appendix A	

4.3 Simulation study

We investigate the performance of the endpoint-corrected estimators defined in (4.4) through simulations using R package `simsalapar` by Hofert and Maechler (2016). The design is as follows: (i) dimension $d\in\{2,4,10\}$; (ii) sample size $n\in\{200,500,1000\}$; (iii) Archimedean generator from the Clayton, Gumbel, Frank and Joe families (see, for example, Nelsen (2006)); (iv) stdf from the following families: Logistic (**LG**), scaled negative extremal Dirichlet (**NSD**) of Belzile and Nešlehová (2017), and discrete spectral measure (**DSM**) of Fougères et al. (2013). The definition of these models may be found in Table 4.1.

The parameters of the Archimedean generator and the stdf were chosen as to cover various scenarios in terms of association, lower/upper tail dependence, and asymmetry. We also intentionally challenge Conditions 3.1–3.3 to explore the robustness of the convergence results. For the sake of brevity, we present the main conclusions of this simulation study and provide representative illustrations; the complete results are available in Appendix A. To evaluate the performance of the estimators, the integrated squared error (ISE) and integrated relative absolute error (IRAE) defined below were used.

$$\begin{aligned}\mathrm{ISE}(A_n)&=\frac{1}{|\Delta_d|}\int_{\Delta_d}\{A_n(\boldsymbol{w})-A(\boldsymbol{w})\}^2\,d\boldsymbol{w},\\ \mathrm{IRAE}(A_n)&=\frac{1}{|\Delta_d|}\int_{\Delta_d}\frac{|A_n(\boldsymbol{w})-A(\boldsymbol{w})|}{A(\boldsymbol{w})}d\boldsymbol{w}.\end{aligned}\tag{4.19}$$

ISE and IRAE were computed using Monte Carlo integration with 10,000 uniformly distributed samples on Δ_d. For each scenario, 1000 Monte Carlo replicates were deemed sufficient to capture the behavior of ISE and IRAE.

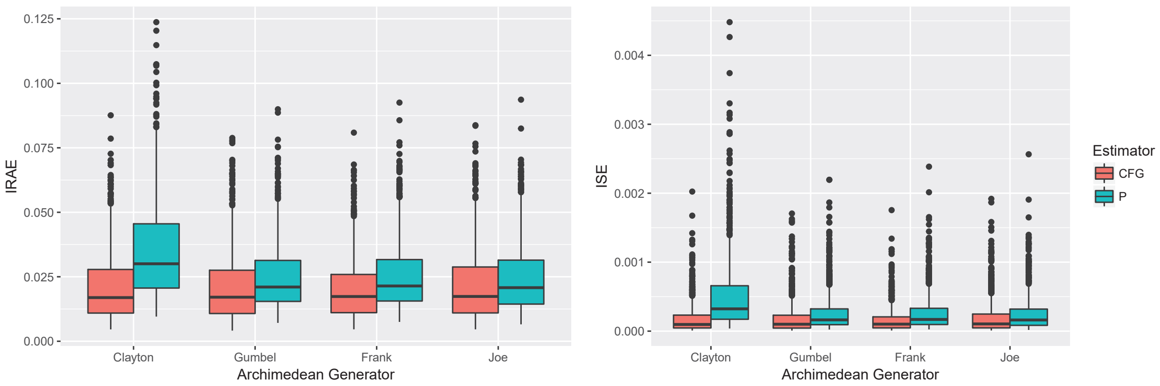

Figure 4.1: Boxplots of IRAE($A_{n,c}$) (left) and ISE($A_{n,c}$) (right) for the Pickands (blue) and CFG (red) type estimators for $n = 200$, $d = 4$, various Archimedean generators with $\tau(\psi) = 1/5$ and the **NSD** stdf with parameters $\boldsymbol{\alpha} = (1, 2, 3, 4)$, $\rho = 0.59$.

Additionally, the finite-sample behavior of the estimators is compared to that of the asymptotic limits obtained in Section 4.1. Observe that from Theorems 4.1–4.2, $\operatorname{var} \mathbb{A}^{\mathrm{CFG}}(\boldsymbol{w})$ and $\operatorname{var} \mathbb{A}^{\mathrm{P}}(\boldsymbol{w})$ are respectively given by

$$\{A(\boldsymbol{w})\}^2 \int_0^1 \int_0^1 \operatorname{cov}\left(\mathbb{C}[\psi\{-\boldsymbol{w}\log(u)\}], \mathbb{C}[\psi\{-\boldsymbol{w}\log(v)\}]\right) \frac{du}{u\log u} \frac{dv}{v\log v},$$

$$\frac{\{A(\boldsymbol{w})\}^4}{\{\mathrm{E}(Z)\}^2} \int_0^1 \int_0^1 \operatorname{cov}\left(\mathbb{C}[\psi\{-\boldsymbol{w}\log(u)\}], \mathbb{C}[\psi\{-\boldsymbol{w}\log(v)\}]\right) \frac{du}{u} \frac{dv}{v},$$

whenever $\boldsymbol{w} \in \mathring{\Delta}_d$. Plots of these asymptotic variances are provided in Figures 4.2 and and corroborate the conclusions drawn from the simulations. They are shown for $d = 2$ as functions of $w \in (0, 1)$, where $\boldsymbol{w} = (w, 1 - w)$.

4.3.1 Comparisons between the Pickands and the CFG-type estimators

We first compared the Pickands and the CFG-type estimators in various scenarios; the results are reported in Tables A1–A6 in Appendix A. Figure 4.1 is representative of the overall pattern, namely that the CFG-type estimator performs better on average both in terms of ISE and IRAE. The superiority of the CFG-type estimator is further supported by Figure 4.2, which shows that in the bivariate case, $\operatorname{var} \mathbb{A}^{\mathrm{CFG}}(w, 1 - w)$ is smaller than $\operatorname{var} \mathbb{A}^{\mathrm{P}}(w, 1 - w)$ for any $w \in (0, 1)$. This is in agreement with Genest and Segers (2009), who observed a similar behavior of the asymptotic variance of the CFG and the Pickands estimator in the bivariate case. In higher dimensions however, the Pickands estimator can sometimes outperform the CFG estimator, although the differences in IRAE and ISE are small; see, e.g., Table A.5 for $d = 10$, small values of $\tau(\psi)$, and the Frank, Gumbel and Joe generators. Figure 4.1 also shows that IRAE is more revealing than ISE, and we concentrate on the former henceforth.

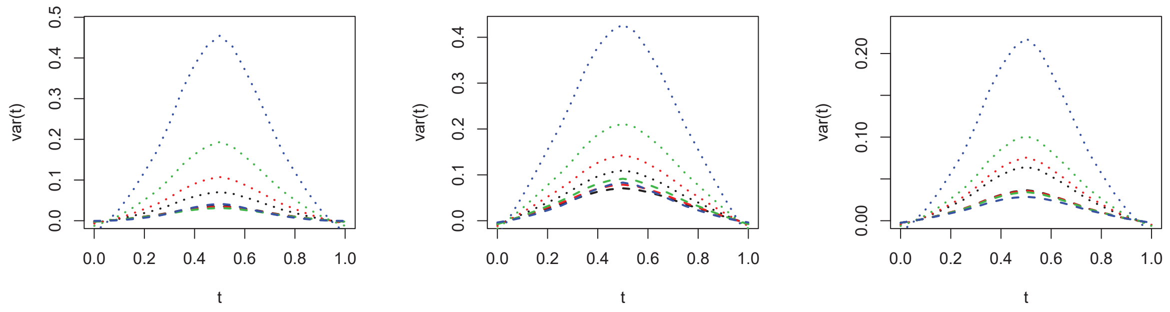

Figure 4.2: Plots of var $\mathbb{A}^{\mathrm{CFG}}(t)$ (dashed) and var $\mathbb{A}^{\mathrm{P}}(t)$ (dotted) for bivariate Archimax copulas with **LG** stdf with parameter $\varrho = 2$. Left: Clayton generator ψ_θ with $\theta = 1/s$ for values of s equal to 5 (black), 5/2 (red), 5/3 (green), 5/4 (blue). Middle: Joe generator ψ_θ with values of $\theta = m$ equal to 1.44 (black), 2.22 (red), 3.83 (green), 8.77 (blue). Right: Frank generator ψ_θ for values of $\tau(\psi)$ equal to 1/5 (black), 2/5 (red), 3/5 (green), 4/5 (blue).

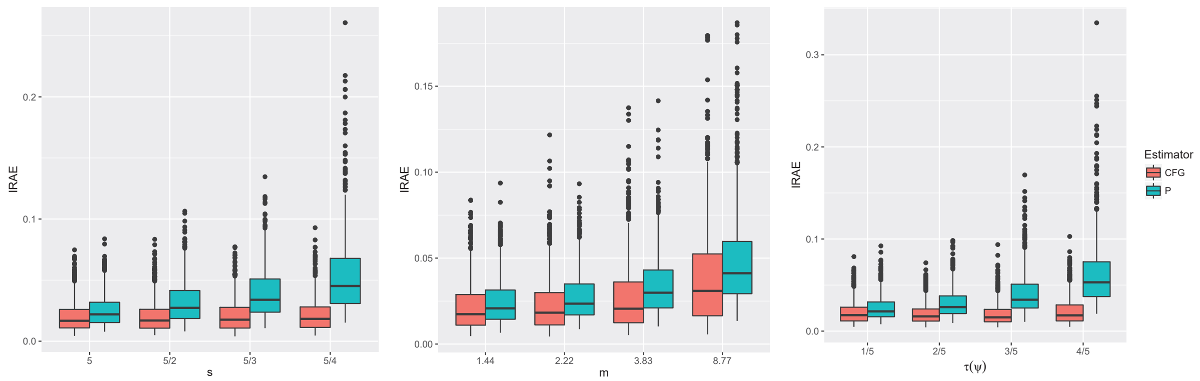

Figure 4.3: Boxplots of IRAE for the Pickands (blue) and CFG (red) estimators for $n = 200$, $d = 4$, and the Clayton generator ψ with $\theta = 1/s$ for various values of s (left), the Joe generator for various values of $\theta = m$ (middle) and Frank for various values of $\tau(\psi) = 1 - (4/\theta)\{1 - D_1(\theta)\}$ (right), where D_1 denotes the Debye function. The stdf is **NSD** with $\boldsymbol{\alpha} = (1, 2, 3, 4)$, $\rho = 0.59$.

Given that the behavior of ψ at zero and infinity played a key role in the conditions of Theorems 4.1 and 4.2, we next investigate the impact of the index of regular variation of ψ and $1 - \psi(1/\cdot)$. Figure 4.3 shows the performance of the estimators for the **NSD** stdf with parameters $\boldsymbol{\alpha} = (1, 2, 3, 4)$, $\rho = 0.59$. In the left panel, the generator is Clayton with parameter θ; the latter satisfies Condition 3.2 (a) with $s = 1/\theta$. This plot reveals that decreasing s has a detrimental effect on $A^{\mathrm{P}}_{n,c}$ while $A^{\mathrm{CFG}}_{n,c}$ is hardly affected. When $s \le 2$, conditions of Theorem 4.2 are no longer met; it is therefore not surprising that the behavior of $A^{\mathrm{P}}_{n,c}$ deteriorates quickly as $s \to 0$. The middle panel of Figure 4.3 explores the effect of m when the generator is Joe, which satisfies Condition 3.1 with $\theta = m$. One can again see that $A^{\mathrm{P}}_{n,c}$ performs worse than $A^{\mathrm{CFG}}_{n,c}$, but this time, increasing m has a negative effect on both estimators. Finally, the right panel of Figure 4.3 shows the effect of dependence of the Archimedean copula C_ψ with generator ψ measured by $\tau(\psi)$, Kendall's tau of the bivariate Archimedean copula with generator ψ, for the Frank generator. In this case, $m = 1$, and increasing $\tau(\psi)$ negatively affects both estimators, although $A_{n,c}^{\mathrm{CFG}}$ is less sensitive. From Figure 4.2, the same conclusions can be drawn about the asymptotic variances.

4.3.2 The effect of the sample size, dimension, and dependence

Given that the CFG-type estimator performed consistently better than $A_{n,c}^{\mathrm{P}}$, we concentrate on the former hereafter and explore the effect of sample size, dimension and dependence. We choose the stdf to be either **LG** with parameter $\varrho = 2$ (all dimensions) or **NSD** with parameters $\boldsymbol{\alpha} = (1, 2)$, $\rho = 0.59$ (for $d = 2$), $\boldsymbol{\alpha} = (1, 2, 3, 4)$, $\rho = 0.59$ (for $d = 4$) and $\boldsymbol{\alpha} = (1, 1, 1, 1, 2, 2, 2, 3, 3, 4)$, $\rho = 0.69$ (for $d = 10$). These parameters are chosen so that the average of pairwise Kendall's taus (see Definition 2.2) of the corresponding d-variate extreme-value copula C_A is 1/2. The Archimedean generator is chosen to be Gumbel with $\theta = 5/3$, which corresponds to Kendall's tau of 2/5 of the corresponding bivariate Archimedean copula C_ψ. The left panel in Figure 4.4 shows the IRAE for various sample sizes when $d = 4$. It is clear that the performance of $A_{n,c}^{\mathrm{CFG}}$ improves with sample size, but also that it depends on the stdf; the CFG-type estimator performs worse when A is **LG**. Other dimensions and Archimedean generators led to the same conclusions. It is worth noting that the asymmetric stdf **NSD** does not lead to better or worse results overall.

The right panel of Figure 4.4 shows the effect of dimension. Unsurprisingly, the performance of $A_{n,c}^{\mathrm{CFG}}$ deteriorates with d. The choice of A has an effect; the latter is most pronounced when $d = 4$, although this may be merely due to the choice of parameters. Again, the same pattern was observed for other sample sizes and Archimedean generators. We also tried the **DSM** Pickands dependence function, which does not satisfy Condition 3.3, because it is not differentiable everywhere. The performance of the CFG-type estimator remained essentially unaffected by this choice of A; see Tables A7–A9 in Appendix A. This is comforting, because Condition 3.3 is virtually impossible to verify from data.

Our next aim was to study the effect of dependence. We restricted ourselves to the **LG** Pickands dependence function; in that case, $C_{\psi,A}$ is exchangeable and measuring dependence can be reduced to the bivariate setting. The first study we conducted focused on Kendall's tau. For a bivariate Archimax copula $C_{\psi,A}$, let $\tau_{\psi,A}$ denote its Kendall's tau $\tau(C_{\psi,A})$; let also $\tau(A) = \tau(C_A)$ and $\tau(\psi) = \tau(C_\psi)$ denote Kendall's tau of the corresponding bivariate extreme-value and Archimedean copula, respectively. From Capéraà et al. (1997),

$$\tau_{\psi,A} = \tau(\psi) + \tau(A) - \tau(\psi)\tau(A). \tag{4.20}$$

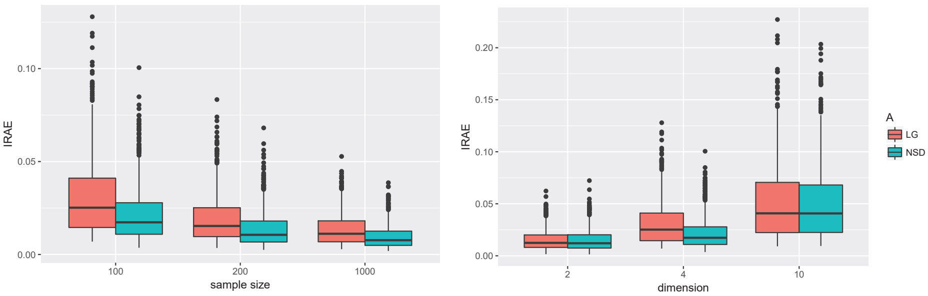

Figure 4.4: Boxplots of IRAE of $A_{n,c}^{\mathrm{CFG}}$ when $d = 4$ and $n \in \{200, 500, 1000\}$ (left), and when $d \in \{2, 4, 10\}$ and $n = 200$ (right). The Pickands dependence functions are **LG** (red) and **NSD** (blue) with coefficient of agreement 1/2; the Archimedean generator is Gumbel with $\theta = 5/3$.

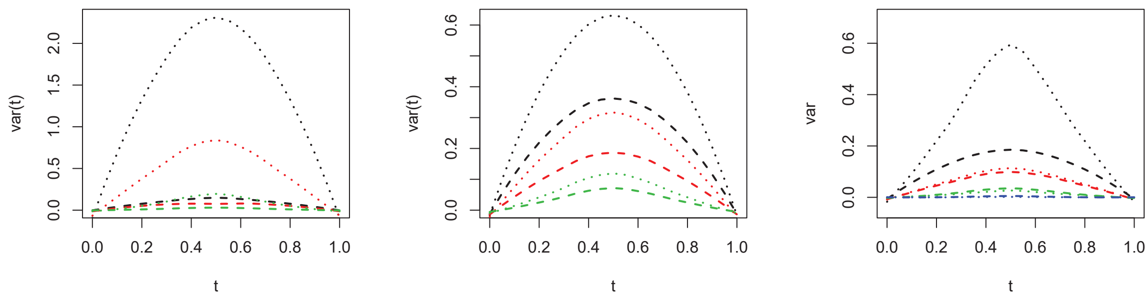

Figure 4.5: Plots of $\operatorname{var} \mathbb{A}^{\mathrm{CFG}}(t)$ (dashed) and $\operatorname{var} \mathbb{A}^{\mathrm{P}}(t)$ (dotted) for bivariate Archimax copulas with stdf **LG**. Left: Joe generator ψ, $\tau(A)$ set for values of 1/5 (black), 2/5 (red), 3/5 (green) and fixed $\tau(\psi, A) = 0.84$. Middle: Frank generator ψ, values of $\lambda_U(A)$ equal to 1/5 (black), 2/5 (red), 3/5 (green) and fixed $\lambda_U(\psi, A) = 0.6$. Right: Clayton generator ψ, values of $\eta(A)$ equal to 0.57 (black), 0.66 (red), 0.76 (green), 0.87 (blue) and fixed $\lambda_L(\psi, A) = 0.4$.

The left panel in Figure 4.6 shows the IRAE of the CFG-type estimator for various values of $\tau_{\psi,A}$ and $\tau(A)$ when $n = 200$ and $d = 10$. The observed trend is that for a fixed $\tau_{\psi,A}$, an increase in $\tau(A)$, which implies a decrease in $\tau(\psi)$, results in lower IRAE. This is corroborated in the asymptotic setting by the left panel of Figure 4.5. There is also a performance gain as $\tau_{\psi,A}$ increases. Conclusions for other Archimedean generators, dimensions and sample sizes are the same; see Tables A10–A12 in Appendix A. The second study focused on the effect of upper tail dependence as measured by λ_U in (2.2). For a bivariate Archimax copula $C_{\psi,A}$ whose generator ψ satisfies Condition 3.1, $\lambda_U(C_{\psi,A}) = 2 - \{2A(1/2)\}^{1/m}$. In the middle panel of Figure 4.6, the stdf is again **LG** with parameter ϱ, so that $A(1/2) = 2^{1/\varrho - 1}$, and the Archimedean generator is Joe with parameter $\theta = m$. Consequently, various values of $\lambda_U(C_{\psi,A})$ can be obtained by varying ϱ and θ. There is a noticeable decrease in IRAE when the contribution of A to $\lambda_U(C_{\psi,A})$ increases, and a slight increase in error for a fixed θ when $\lambda_U(C_{\psi,A})$ increases. A similar conclusion can be drawn in terms of the asymptotic variances from Figure 4.5 (middle panel). The same

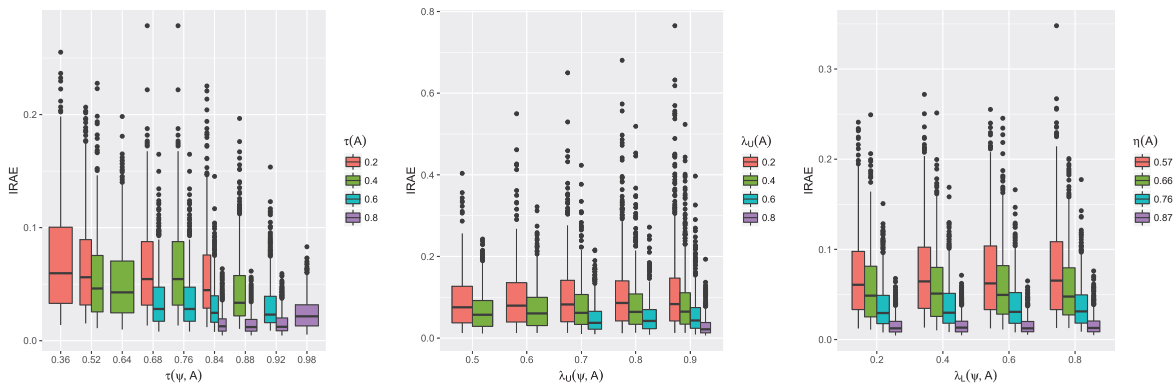

Figure 4.6: Boxplots of IRAE of $A_{n,c}^{\mathrm{CFG}}$ when $n = 200$, $d = 10$ and the Pickands dependence function is **LG** for all panels. The Archimedean generators are Frank (left), Joe (middle) and Clayton (right). In the right panel, $\eta_L(A) = 1/\{2A(1/2)\} = 2^{-1/\rho}$ is the lower tail dependence index of Ledford and Tawn (1996).

pattern was observed for other choices of n and d; see Table A13 in Appendix A.

The last study focused on the effect of lower tail dependence as measured by λ_L in (2.3). For a bivariate Archimax copula $C_{\psi,A}$ whose generator ψ satisfies Condition 3.2 (a), $\lambda_L(C_{\psi,A}) = \{2A(1/2)\}^{-s}$. Again, we considered the **LG** Pickands dependence function. As the Archimedean generator we choose the Clayton generator, which is such that $s = 1/\theta$. The right panel of Figure 4.6 shows that the effects of lower and upper tail dependence are similar: an increase in the contribution of A to λ_L leads to lower IRAE. This agrees with the right panel of Figure 4.5. There is also a slight decrease in performance when θ is fixed and $\lambda_L(C_{\psi,A})$ increases. The same pattern occurred for other choices of n and d; see Table A14 in Appendix A.

Chapter 5

Semiparametric inference for Archimax copulas

Chapter 4 focused on the nonparametric estimation of the stable tail dependence function under the assumption that the distortion function ψ is known. Building upon these results, we can now relax this assumption by supposing instead that ψ belongs to a parametric family, i.e. $\psi \in \Psi = \{\psi_\theta, \theta \in \mathcal{O}\}$, $\mathcal{O} \subset \mathbb{R}^p$. The Archimedean copula family has a very rich literature surrounding it with many parametric families having been studied extensively studied. Their flexibility translates well into modeling with Archimax copulas since generators can be chosen to capture certain aspects of the data at hand. For example, if the dataset exhibits lower tail dependence, the Clayton generator could potentially be a good candidate. Once the parametric family is chosen, θ needs to be estimated without the knowledge of ℓ, and we present an idea on how to do this for one-parameter families in Section 5.1. Section 5.2 contains the estimators of the stable tail dependence function, which are adapted from those of Chapter 4. Section 5.3 gathers the conditions on the parametric family for ψ needed in order to study the convergence of the estimators as is done in Section 5.4. Finally, Section 5.5 contains the proofs of said convergence results.

5.1 Estimation of ψ

How ψ can be estimated without the knowledge of ℓ, again assuming that $\psi \in \Psi$ where $\Psi = \{\psi_\theta, \theta \in \mathcal{O}\}$? Recall that under the assumptions of Proposition 3.1, θ and ℓ are then identifiable. In this section, we propose a simple moment-based procedure for the most common scenario where $\mathcal{O} \subseteq \mathbb{R}$.

First consider an arbitrary bivariate copula C and a pair $(U_1, U_2) \sim C$. The distribution function K_C of the random variable $W_C = C(U_1, U_2)$ is called the Kendall distribution, see Barbe et al. (1996). If $C = C_{\psi,A}$ is Archimax, it is known from Eq. (13) in Capéraà et al. (2000) that for any $w \in [0, 1]$, $K_{C_{\psi,A}}(w) = K_{C_\psi}(w) + \phi(w)/\phi'(w)\tau(A)$, where $\tau(A)$ is Kendall's tau of C_A. Hence for any $k \in \mathbb{N}$, the kth moment of $W_{C_{\psi,A}}$

satisfies

$$m_k = \mathrm{E}(W^k_{C_{\psi,A}}) = \tau(A)\,\frac{1}{k+1} + \{1-\tau(A)\}\,\mathrm{E}(W^k_{C_\psi})\,. \tag{5.1}$$

Equations (5.1) for $k = 1$ and $k = 2$ then lead to the following identity:

$$\frac{1-2\,\mathrm{E}(W_{C_\psi})}{1-3\,\mathrm{E}(W^2_{C_\psi})} = \frac{1-2m_1}{1-3m_2}. \tag{5.2}$$

The left-hand side depends only on the Archimedean generator and is thus a function of θ, say f. Assuming that ψ is twice differentiable, Theorem 4.3.4 in Nelsen (2006) and partial integration yield that for any $\theta \in \mathcal{O}$,

$$f(\theta) = \frac{1-2\,\mathrm{E}(W_{C_{\psi_\theta}})}{1-3\,\mathrm{E}(W^2_{C_{\psi_\theta}})} = \frac{\int_0^{x_{\psi_\theta}} x\{\psi'_\theta(x)\}^2 dx}{3\int_0^{x_{\psi_\theta}} x\psi_\theta(x)\{\psi'_\theta(x)\}^2 dx}. \tag{5.3}$$

The following example provides explicit expressions for f for three families of generators; in each case, f is strictly monotone in θ.

Example 5.1. *For the Clayton generator given in Table 4.1,* $\mathrm{E}(W^k_{\psi_\theta}) = (\theta+1)/\{(k+1)(\theta+k+1)\}$ *for any* $k \in \mathbb{N}$. *Consequently,*

$$f(\theta) = \theta + 3/\{2(\theta+2)\}.$$

Next, consider the Genest–Ghoudi family Genest and Ghoudi (1994) whose generator is given, for any $x \in [0,1]$, *by* $\psi_\theta(x) = (1-x^\theta)^{1/\theta}$ *for* $\theta \in (0,1]$. *Here,* $\mathrm{E}(W^k_{\psi_\theta}) = (1-\theta)/(k+1-\theta)$, *for any* $k \in \mathbb{N}$. *Hence,*

$$f(\theta) = 3 - \theta/(4-2\theta).$$

Finally, consider the Frank generator given in Table 4.1. For $j \in \mathbb{N}$, *let* $D_j(\theta) = (j/\theta^j)\int_0^\theta t^j/(e^t-1)dt$ *denote the Debye function (Abramowitz and Stegun, 1964, Chap. 27). Here,* (5.3) *yields that for any* $\theta \in \mathbb{R}$,

$$f(\theta) = \frac{4\theta - 4\theta D_1(\theta)}{3\{2\theta - \theta D_2(\theta) + 4D_1(\theta) - 4\}}.$$

If f is one-to-one, as was the case in Example 5.1, Eq. (5.2) can be used to construct an estimator of θ. Following Ben Ghorbal et al. (2009), let $I_{ij} = \mathbf{1}(X_i \le X_j, Y_i \le Y_j)$ for all $i,j \in \{1,\dots,n\}$ and set

$$m_{n,1} = \frac{1}{n(n-1)}\sum_{i\neq j} I_{ij}, \quad m_{n,2} = \frac{1}{n(n-1)(n-2)}\sum_{i\neq j\neq k} I_{ij}I_{kj}.$$

As $m_{n,1}$ and $m_{n,2}$ are U-statistics with square integrable kernels, the results of these authors imply that $\sqrt{n}\,\{(m_{n,1},m_{n,2}) - (\mathrm{E}(W_C),\mathrm{E}(W_C^2))\} \rightsquigarrow \mathcal{N}(\mathbf{0},\Sigma)$ as $n \to \infty$; the entries of Σ are given in Proposition 2 therein.

Next, provided f has an inverse $f^{\leftarrow}$, define $h : \mathbb{R}^2 \to \mathbb{R}$ by

$$h(m_1, m_2) = f^{\leftarrow}\left(\frac{1-2m_1}{1-3m_2}\right)$$

and set $\theta_n = h(m_{n,1}, m_{n,2})$. Assuming h has continuous partial derivatives that are non-zero at (m_1, m_2) and using the delta method, one gets that $\sqrt{n}\,(\theta_n - \theta) \rightsquigarrow \mathcal{N}[0, J_h(m_1, m_2)\,\Sigma J_h(m_1, m_2)^\top]$, where J_h is the 2×1 Jacobian matrix of h. Consistent plug-in estimators of the entries of Σ are provided in Ben Ghorbal et al. (2009). For small n, the calculations presented in that paper can also be used to compute and estimate the finite-sample variance-covariance matrix of $(m_{n,1}, m_{n,2})$.

Example 5.2. *For the Clayton family, $\theta_n = S_n/R_n$, where*

$$S_n = 8m_{n,1} - 9m_{n,2} - 1, \quad R_n = 1 - 4m_{n,1} + 3m_{n,2}. \tag{5.4}$$

Then $\sqrt{n}\,(\theta_n - \theta) = \sqrt{n}\,\{h(m_{n,1}, m_{n,2}) - h(m_1, m_2)\} \rightsquigarrow \mathcal{N}(0, \sigma^2)$, where σ^2 is defined as follows as a function of $S = 8m_1 - 9m_2 - 1$ and $R = 1 - 4m_1 + 3m_2$:

$$\begin{aligned}\sigma^2 = \frac{1}{R^4}\big\{&R^2(64\Sigma_{11} + 81\Sigma_{22} - 144\Sigma_{12}) \\ &+ S^2(16\Sigma_{11} + 9\Sigma_{22} - 24\Sigma_{12}) - 2RS(32\Sigma_{11} - 27\Sigma_{22} + 50\Sigma_{12})\big\}.\end{aligned} \tag{5.5}$$

Note that the numerator S_n in (5.4) is the quantity on which the test for bivariate extreme-value dependence of Ghoudi et al. (1998) is based. These authors showed that when C is an extreme-value copula, $8\,E(W_C) - 9\,\mathrm{E}(W_C^2) - 1 = 0$. When $\theta = 0$, the Clayton generator becomes $\psi(t) = e^{-t}$ and $C_{\psi,A} = C_A$ is an extreme-value copula.

For the Genest–Ghoudi family, $\theta_n = -S_n/R_n$, where S_n and R_n are as in (5.4). Hence $\sqrt{n}\,(\theta_n - \theta) \rightsquigarrow \mathcal{N}(0, \sigma^2)$, where σ^2 is given by (5.5).

For the bivariate Frank family, the function f is one-to-one but its inverse is not explicit. Therefore, both the estimator and the asymptotic variance are not explicit either. An estimate of θ can be obtained numerically and its asymptotic variance can be studied via resampling.

In the multivariate case, a generalization of (5.1) does not seem possible. We thus propose to use $\theta_n = 2\sum_{j<k} \theta_{n,jk}/\{d(d-1)\}$, where $\theta_{n,jk}$ is the above moment-based estimator of θ based on the bivariate sample $(X_{1j}, X_{1k}), \ldots, (X_{nj}, X_{nk})$. A heuristic approach for checking whether averaging the pair-wise estimates is reasonable is presented in the next section.

5.2 Estimation of ℓ when ψ is unknown

We now focus on the nonparametric estimator of A and its asymptotic properties assuming that an estimator of θ is available. Once θ has been estimated by θ_n in such a way that

$\theta_n \in \mathcal{O}$ for all $n \in \mathbb{N}$, the Pickands or CFG-type estimators of A can be constructed as in Chapter 4 with ψ replaced by ψ_{θ_n}. For every $\boldsymbol{w} \in \Delta_d$, and $i \in \{1, \ldots, n\}$, let

$$\hat{\xi}_{i,n}(\boldsymbol{w}) = \min\{\phi_{\theta_n}(\hat{U}_{ij})/w_1, \ldots, \phi_{\theta_n}(\hat{U}_{ij})/w_d\}$$

with the convention that $\phi_{\theta_n}(\hat{U}_{ij})/w_j = \infty$ when $w_j = 0$. As before, $\hat{\xi}_{i,n}(\boldsymbol{w})$ is finite for every $i \in \{1, \ldots, n\}$. When $\mathrm{E}(\log Z)$ and $\mathrm{E}(Z)$ exist, respectively, the CFG and Pickands-type estimators are given, for each $\boldsymbol{w} \in \Delta_d$, by

$$\log \hat{A}_n^{\mathrm{CFG}}(\boldsymbol{w}) = \mathrm{E} \log Z - \frac{1}{n} \sum_{i=1}^n \log \hat{\xi}_{i,n}(\boldsymbol{w}), \quad \hat{A}_n^{\mathrm{P}}(\boldsymbol{w}) = n\,\mathrm{E}(Z) \Big/ \sum_{i=1}^n \hat{\xi}_{i,n}(\boldsymbol{w}).$$

Because ψ is estimated by ψ_{θ_n} rather than fixed, the weak limit of

$$\hat{\mathbb{A}}_n^{\mathrm{CFG}} = \sqrt{n}\,(\hat{A}_n^{\mathrm{CFG}} - A), \quad \hat{\mathbb{A}}_n^{\mathrm{P}} = \sqrt{n}\,(\hat{A}_n^{\mathrm{P}} - A) \tag{5.6}$$

is no longer the process given in Theorems 4.1 and 4.2, respectively.

5.3 Regularity conditions

The conditions on the parametric family $\Psi = \{\psi_\theta, \theta \in \mathcal{O}\}$ are considered. In what follows, $\|\cdot\|_2$ denotes the ℓ_2-norm and $\mathring{\mathcal{O}}$ denotes the interior of $\mathcal{O}$.

Condition 5.1. *For all $\theta \in \mathcal{O}$, $\phi_\theta(0) = x_{\psi_\theta}$ is identical and equal to x_Ψ.*

Condition 5.2. *Let $\Theta_n = \sqrt{n}\,(\theta_n - \theta_0)$. Whenever $\theta_0 \in \mathring{\mathcal{O}}$, $n \to \infty$, $(\hat{\mathbb{C}}_n, \Theta_n) \rightsquigarrow (\mathbb{C}, \Theta)$ in $\ell^\infty([0,1]^d) \times \mathbb{R}^p$ and the limit is centered Gaussian.*

Condition 5.3. *For any $\theta \in \mathring{\mathcal{O}}$, the gradient*

$$\dot{\psi}_\theta(t) = (\dot{\psi}_{\theta,1}(t), \ldots, \dot{\psi}_{\theta,p}(t))^\top = (\partial\psi_\theta(t)/\partial\theta_1, \ldots, \partial\psi_\theta(t)/\partial\theta_p)^\top$$

exists and is continuous for all $t \in [0, x_\Psi)$.

The following condition is needed for the CFG-type estimator.

Condition 5.4. *For any $\theta \in \mathring{\mathcal{O}}$, there exists an $\omega \in (0, 1/2)$ and a bounded, non-negative function h_θ on $[0, x_\Psi)$ such that for each $j \in \{1, \ldots, p\}$, $|\dot{\psi}_{\theta,j}|/h_\theta$ is bounded on $[0, x_\Psi)$,*

$$\int_0^{x_\Psi} \frac{h_\theta^\omega(t)}{t} dt < \infty, \quad \int_0^{x_\Psi} \frac{h_\theta(t)}{t} dt < \infty,$$

and such that $\Upsilon_\theta(\epsilon) \to 0$ for $\epsilon \to 0$, where for any $\epsilon > 0$,

$$\Upsilon_\theta(\epsilon) = \sup_{\theta' \in \mathcal{O}, \|\theta' - \theta\|_2 \leq \epsilon} \; \sup_{t \in [0, x_\Psi)} \frac{\|\dot{\psi}_{\theta'}(t) - \dot{\psi}_\theta(t)\|_2}{h_\theta(t)}.$$

The following condition pertains to the Pickands-type estimator.

Condition 5.5. *For any $\theta \in \mathring{\mathcal{O}}$, there exists an $\omega \in (0, 1/2)$ and a bounded, non-negative function h_θ on $[0, x_\Psi)$ such that for each $j \in \{1, \ldots, p\}$, $|\dot{\psi}_{\theta,j}|/h_\theta$ is bounded on $[0, x_\Psi)$,*

$$\int_0^{x_\Psi} h_\theta^\omega(t)dt < \infty, \qquad \int_0^{x_\Psi} h_\theta(t)dt < \infty,$$

and such that $\Upsilon_\theta(\epsilon) \to 0$ for $\epsilon \to 0$, where $\Upsilon_\theta(\epsilon)$ is as in Condition 5.4.

Finally, two more conditions are needed, each assuming Condition 5.3.

Condition 5.6. *For any $\theta \in \mathring{\mathcal{O}}$, the Hessian $\ddot{\psi}_\theta(t) = (\ddot{\psi}_{\theta,jk}(t))_{j,k} = (\partial^2 \psi_\theta(t)/\partial\theta_j\partial\theta_k)_{j,k}$ exists and is continuous for all $t \in [0, x_\Psi)$. Furthermore, for each $j, k \in \{1, \ldots, p\}$, $\ddot{\psi}_{\theta,jk}(t) \to 0$ as $t \to 0$ and as $t \to x_\Psi$, and*

$$\lim_{\epsilon \downarrow 0} \sup_{\theta' \in \mathcal{O}, \|\theta' - \theta\|_2 \le \epsilon} \sup_{t \in [0, x_\Psi)} \|\ddot{\psi}_{\theta'}(t) - \ddot{\psi}_\theta(t)\|_{\mathrm{E}} = 0,$$

where $\|\cdot\|_{\mathrm{E}}$ denotes the entrywise 1-norm, i.e., $\|A\|_{\mathrm{E}} = \sum_{j,k} |A_{jk}|$.

Condition 5.7. *For each $j \in \{1, \ldots, p\}$, $\theta \in \mathring{\mathcal{O}}$ and any $\delta > 0$ such that $\{\theta' \in \mathbb{R}^p : \|\theta - \theta'\| < \delta\} \subset \mathring{\mathcal{O}}$,*

$$\lim_{u \downarrow 0} \sup_{\theta' : \|\theta - \theta'\|_2 < \delta} \frac{\dot{\psi}_{\theta',j}\{\phi_{\theta'}(u)\}}{\sqrt{u}} = \lim_{u \downarrow 0} \sup_{\theta' : \|\theta - \theta'\|_2 < \delta} \frac{\dot{\psi}_{\theta',j}\{\phi_{\theta'}(1-u)\}}{\sqrt{u}} = 0.$$

In the following, the above conditions are validated for the Clayton family of Archimedean generators.

Example 5.3 (Verification of the regularity conditions for the Clayton family)**.** *Consider the Clayton family with generator given, for any $x \ge 0$, by $\psi_\theta(x) = (1 + \theta x)^{-1/\theta}$ where $\theta \in \mathcal{O} = [0, \infty)$; when $\theta = 0$, $\psi_\theta(x) = e^{-x}$. For this family, θ may be estimated for example as in Example 5.2; to make the estimator intrinsic, one can use $\theta_n^* = \max(\theta_n, 0)$. Because θ_n is consistent, $|\theta_n^* - \theta_n| = o_P(1)$ whenever the true parameter value θ_0 is strictly positive. Thus for $\theta_0 > 0$, $\sqrt{n}(\theta_n^* - \theta_0)$ is asymptotically centered Gaussian, with the same variance σ^2 as given in Example 5.2.*

Condition 5.1. For this family, for any $\theta \ge 0$, $\phi_\theta(0) = x_\Psi = \infty$.

Condition 5.2. The validity of this condition follows from the joint convergence of ($\hat{\mathbb{C}}_n$, $\sqrt{n}\{(m_{n,1}, m_{n,2}) - (\mathrm{E}(W_C), \mathrm{E}(W_C^2))\}$). Because $m_{n,1}$ and $m_{n,2}$ are U-statistics with squared-integrable kernels, the latter can be established using Hájek's projection technique; see van der Vaart (1998), for example.

Condition 5.3. For all $\theta \in (0, \infty)$ and $x \in (0, \infty)$, $\dot{\psi}_\theta$ exists and is continuous for all $x \in [0, \infty)$. In fact,

$$\dot{\psi}_\theta(x) = \frac{1}{\theta^2}(1 + \theta x)^{-1/\theta}\left\{\ln(1 + \theta x) - \frac{\theta x}{1 + \theta x}\right\} = \frac{1}{\theta^2}(1 + \theta x)^{-1/\theta - 1}\int_1^{1+\theta x} \ln(t)dt$$

given that the derivative of $\{x\ln(x) - x + 1\}$ *is* $\ln(x)$.

Condition 5.4. An admissible function h_θ *is defined for any* $x \geq 0$ *by*

$$h_\theta(x) = h(x;\theta,\delta,\eta) = x^\delta(1+\theta x)^{-\eta} \,,$$

where $0 < \delta < \eta$ *satisfy* $0 < \delta < 1$ *and* $\eta - \delta < 1/\theta$. *One can first quickly check that the integrals* $\int_0^{x_\Psi} h_\theta^\omega(t)dt/t$ *and* $\int_0^{x_\Psi} h_\theta(t)dt/t$ *are both finite, as soon as* $\delta > 0$ *and* $\delta < \eta$, *and this for any* $\omega \in (0,1/2)$. *Besides,*

$$\frac{|\dot{\psi}_\theta|}{h_\theta}(x) = x^{-\delta}(1+\theta x)^{\eta-1/\theta}\left\{\frac{1}{\theta^2}\ln(1+\theta x) - \frac{x}{\theta}(1+\theta x)^{-1}\right\} \,,$$

so that it is a bounded function on $[0,\infty)$ *as soon as* $0 < \delta < 1$ *and* $\eta - \delta < 1/\theta$. *The last point to check is then that* $\Upsilon_\theta(\epsilon) \to 0$ *when* $\epsilon \to 0$, *where* $\Upsilon_\theta(\epsilon)$ *is defined for any* $\epsilon > 0$ *in Condition 5.4. As soon as* ψ_θ *is* $\mathcal{C}^2$, *one can write that for any* $x \in [0,\infty)$, *and for any* $\theta' \in \mathcal{O}$ *such that* $|\theta' - \theta| < \epsilon$,

$$\frac{|\dot{\psi}_{\theta'}(x) - \dot{\psi}_\theta(x)|}{h_\theta(x)} \leq \sup_{\theta''\in\mathcal{O},|\theta''-\theta|<\epsilon} |\ddot{\psi}_{\theta''}(x)| \frac{|\theta-\theta'|}{h_\theta(x)} \leq \epsilon \sup_{\theta''\in\mathcal{O},|\theta''-\theta|<\epsilon} \frac{|\ddot{\psi}_{\theta''}(x)|}{h_\theta(x)} \,.$$

Now choose an arbitrary ϵ_0 *such that* $0 \leq \epsilon \leq \epsilon_0$. *One can also write*

$$\frac{|\dot{\psi}_{\theta'}(x) - \dot{\psi}_\theta(x)|}{h_\theta(x)} \leq \epsilon \sup_{\theta''\in\mathcal{O},|\theta''-\theta|<\epsilon_0} \frac{|\ddot{\psi}_{\theta''}(x)|}{h_\theta(x)} \,.$$

For the Clayton generator, one gets for any $\theta > 0$ *and any* $x \in [0,\infty)$,

$$\begin{aligned}\ddot{\psi}_\theta(x) = &\frac{1}{\theta^4}(1+\theta x)^{-1/\theta}\Big\{\ln(1+\theta x) - \frac{\theta x}{1+\theta x}\Big\}^2 \\ &+ \frac{1}{\theta^3}(1+\theta x)^{-1/\theta}\Big\{-2\ln(1+\theta x) + 3\frac{\theta x}{1+\theta x} - \frac{\theta x}{(1+\theta x)^2}\Big\}. \qquad (5.7)\end{aligned}$$

Thus for any $\theta'' \in (\theta - \epsilon_0, \theta + \epsilon_0)$ *and* $x \in [0,\infty)$, $\ddot{\psi}_{\theta''}(x) = \sum_{i=1}^6 g_i(x,\theta'')$ *in terms of six functions* $g_i(x,t) = \alpha_{0,i}t^{\alpha_{1,i}}(1+tx)^{\alpha_{2,i}}\{\ln(1+tx)\}^{\alpha_{3,i}}$, *for fixed reals* $\alpha_{k,i}$, *where* $k = 0,\ldots 3$ *and* $i = 1,\ldots,6$. *Making use of the fact that* $\theta - \epsilon_0 \leq \theta'' \leq \theta + \epsilon_0$, *one can then majorize each of the terms* $|g_i(x,\theta'')|$ *by, say,* $\tilde{M}^i_{\theta,\epsilon_0}(x)$, *and obtain that*

$$\sup_{\theta''\in\mathcal{O},|\theta''-\theta|<\epsilon_0} \frac{|\ddot{\psi}_{\theta''}(x)|}{h_\theta(x)} \leq M_{\theta,\epsilon_0}(x) \,,$$

where $M_{\theta,\epsilon_0}(x)$ *is defined as* $M_{\theta,\epsilon_0}(x) = \sum_{i=1}^6 \tilde{M}^i_{\theta,\epsilon_0}(x)/h_\theta(x)$. *One can then check that when* x *tends to 0,* $M_{\theta,\epsilon_0}(x) = O(x^{1-\delta})$, *which tends to 0 since* $\delta < 1$. *Analogously, when* x *tends to infinity, one gets that* $M_{\theta,\epsilon_0}(x) = O(x^{\eta-\delta-1/\theta})$, *which tends to 0 since* $\eta - \delta < 1/\theta$. *As a consequence,*

$$\bar{M}_{\theta,\epsilon_0} := \sup_{x\in[0,\infty)} M_{\theta,\epsilon_0}(x) < \infty \,.$$

This allows to conclude that

$$\Upsilon_\theta(\epsilon) = \sup_{\theta' \in \mathcal{O}, |\theta'-\theta|<\epsilon} \sup_{x\in[0,x_\Psi)} \frac{|\dot{\psi}_{\theta'}(x) - \dot{\psi}_\theta(x)|}{h_\theta(x)} \le \epsilon\, \bar{M}_{\theta,\epsilon_0}\ ,$$

which tends to 0 as ϵ tends to 0 and leads to the desired result.

Condition 5.5. For the Clayton family, Condition 3.2 (a) holds with $s = 1/\theta$. Because $s > 2$ in Theorem 5.2, $\theta < 1/2$. Furthermore, from the proof of Lemma 5.10, $s\omega > 1$ so that $\omega \in (\theta, 1/2)$. For a fixed ω in this interval, a suitable choice for h_θ is

$$h_\theta(x) = (1+\theta x)^{-\eta},$$

where $\eta \in (1/\omega, 1/\theta)$. One can then easily check that $\int_0^\infty h_\theta(x)dx$ and $\int_0^\infty h_\theta^\omega(x)dx$ are finite and also that $|\dot{\psi}_\theta|/h_\theta$ is bounded on $[0,\infty)$. It remains to verify that $\Upsilon_\theta(\epsilon) \to 0$ as $\epsilon \to 0$. To this end, recall from (5.7) that ψ is twice continuously differentiable w.r.t. θ; it is helpful to note that for all $\theta > 0$ and $x \in [0,\infty)$,

$$\begin{aligned}\ddot{\psi}_\theta(x) = {} & \frac{1}{\theta^4}(1+\theta x)^{-1/\theta-2}\Big\{\int_1^{1+\theta x} \ln(t)dt\Big\}^2 \\ & + \frac{1}{\theta^3}(1+\theta x)^{-1/\theta-1}\Big\{\int_1^{1+\theta x} \ln(t)dt\Big\}\Big\{-2 - \frac{\theta x}{1+\theta x}\Big\} \\ & + \frac{1}{\theta^3}(1+\theta x)^{-1/\theta-1}\ln(1+\theta x)\frac{\theta x}{1+\theta x} \qquad (5.8)\end{aligned}$$

and that for all $x \in [0, x_\Psi)$, $\theta > 0$, and $k \in \mathbb{N}$,

$$(1+\theta x)^{-1/\theta-k}\Big(\int_1^{1+\theta x} \ln(t)dt\Big)^k \le (1+\theta x)^{-1/\theta}\{\ln(1+\theta x)\}^k. \qquad (5.9)$$

Because $\eta < 1/\theta$, there exists some small $\epsilon_0 \in (0,\theta)$ so that $\eta < 1/(\theta+\epsilon_0)$. Given that for any $\epsilon < \epsilon_0$, one has that

$$\Upsilon_\theta(\epsilon) \le \epsilon \sup_{\theta'':|\theta-\theta''|\le\epsilon} \sup_{x\ge 0} \frac{|\ddot{\psi}_{\theta''}(x)|}{h_\theta(x)}$$

and hence it suffices to show that $|\ddot{\psi}_{\theta''}(x)|/h_\theta(x)$ is bounded from above for all $x \ge 0$ and $\theta'' \in (\theta-\epsilon_0, \theta+\epsilon_0)$. From (5.9) and the fact that for any $t \in [1,\infty)$, $k \in \mathbb{N}$ and $\lambda > 0$, $t^{-\lambda}\{\ln(t)\}^k$ is bounded above by $(k/\lambda e)^k$, one has that $|\ddot{\psi}_{\theta''}(x)|/h_\theta(x)$ is bounded above by

$$\Big(\frac{\theta}{\theta-\epsilon_0}\Big)^\eta \Big\{\Big(\frac{1}{\theta-\epsilon_0}\Big)^4\Big(\frac{1}{\theta+\epsilon_0}-\eta\Big)^{-2}\Big(\frac{2}{e}\Big)^2 + 4\Big(\frac{1}{\theta-\epsilon_0}\Big)^3\Big(\frac{1}{\theta+\epsilon_0}-\eta\Big)^{-1}\Big(\frac{1}{e}\Big)\Big\}.$$

Condition 5.6. Clearly, the function in (5.7) is continuous for all x and $\ddot{\psi}_\theta(x) \to 0$ when $x \to 0$ as well as when $x \to \infty$. To verify the smoothness condition of $\ddot{\psi}_\theta$, fix an arbitrary $\epsilon \in (0,\theta)$. It suffices to show that $|\dddot{\psi}_{\theta''}(x)|$ is bounded from above for all $\theta'' \in (\theta-\epsilon, \theta+\epsilon)$ and $x \in [0, x_\Psi)$. Using (5.8) and (5.9), $\dddot{\psi}_{\theta''}(x)$ can be computed to be a sum of finitely *many terms, each of which is, for any $x \in [0, x_\Psi)$, bounded in absolute value from above by a term of the form*

$$\frac{c}{(\theta'')^m}(1+\theta'' x)^{-1/\theta''}\{\ln(1+\theta'' x)\}^k \tag{5.10}$$

for some positive constant c, independent of θ and ϵ, and some $m, k \in \mathbb{N}$. Because for any $t \in [1, \infty)$, $t^{-1/\theta''}\{\ln(t)\}^k$ is bounded above by $(\theta'')^k(k/e)^k$, the term in (5.10) is further bounded above, for $\theta'' \in (\theta-\epsilon, \theta+\epsilon)$, by $\{c/(\theta-\epsilon)^m\}(\theta+\epsilon)^k(k/e)^k$ which converges to $\{c/\theta^m\}\theta^k(k/e)^k$ as $\epsilon \to 0$.

Condition 5.7. Because for all $u \in (0, 1]$, $\phi_\theta(u) = (u^{-\theta} - 1)/\theta$,

$$\dot{\psi}_\theta\{\phi_\theta(u)\} = \frac{u}{\theta^2}\{-\theta\ln(u) - (1-u^\theta)\}$$

Fix an arbitrary $\theta > 0$ and $\delta \in (0, \theta)$. Then for any $\theta' \in (\theta-\delta, \theta+\delta)$,

$$\begin{aligned}\frac{1}{\sqrt{u}}|\dot{\psi}_{\theta'}\{\phi_\theta(u)\}| &\le \frac{\sqrt{u}}{\theta'}\{-\ln(u)\} + \frac{\sqrt{u}}{(\theta')^2}(1-u^{\theta'}) \\ &\le \frac{\sqrt{u}}{\theta-\delta}\{-\ln(u)\} + \frac{\sqrt{u}}{(\theta-\delta)^2}(1-u^{\theta+\delta}).\end{aligned}$$

Clearly, the upper bound converges to 0 as $u \to 0$. Similarly, for any $\theta' \in (\theta-\delta, \theta+\delta)$, $|\dot{\psi}_{\theta'}\{\phi_\theta(1-u)\}|/\sqrt{u}$ is at most

$$\frac{\sqrt{u}(1-u)}{\theta-\delta}\frac{\{-\ln(1-u)\}}{u} + \frac{\sqrt{u}(1-u)}{(\theta-\delta)^2}\frac{1-(1-u)^{\theta+\delta}}{u}.$$

Again, the upper bound converges to 0 as $u \to 0$.

5.4 Asymptotic behavior

Under the conditions elicited in Section 5.3, the following two results may be established. The proofs are rather tedious and may be found in Section 5.5. In the following, Θ denotes the weak limit of $\sqrt{n}\,(\theta_n - \theta_0)$ and $\dot{\psi}_\theta(x)$ is the derivative of $\psi_\theta(x)$ with respect to θ. The existence of the latter for all $x \in [0, x_{\psi_\theta})$ is guaranteed by Condition 5.2; we set $\dot{\psi}_\theta(x) \equiv 0$ for $x \ge x_{\psi_\theta}$ in order to simplify the expression of the limiting process.

Theorem 5.1. *Suppose that $\boldsymbol{X}_1, \boldsymbol{X}_2, \ldots$ is a stationary, alpha-mixing sequence with $\alpha^{[\boldsymbol{X}]}(k) = O(a^k)$, as $k \to \infty$, for some $a \in (0, 1)$. Suppose that the marginals of the stationary distribution are continuous and the corresponding copula belongs to the class of d-variate Archimax copulas $\mathcal{C}_\Psi$ whose stdfs are arbitrary with $\ell \ne \ell_M$ and whose Archimedean generators belong to a parametric family $\Psi = \{\psi_\theta, \theta \in \mathcal{O}\}$, $\mathcal{O} \subseteq \mathbb{R}^p$. Assume that $\mathcal{C}_\Psi$ satisfies the conditions of Proposition 3.1. Suppose further that the true parameter value θ_0 is in the interior $\mathring{\mathcal{O}}$ of $\mathcal{O}$, that ψ_{θ_0} is q-monotone for some $q \ge 3$ and such that ψ''_{θ_0} exists and is continuous on $(0, \infty)$. Further assume that ψ_{θ_0} satisfies Conditions 3.1 and 3.3, as well as either Condition 3.2 (a) or Condition 3.2 (b) with the additional* *requirement that* $-\log(\psi_{\theta_0})$ *is concave on* $(0, x_{\psi_{\theta_0}})$*. Finally, assume that Conditions 5.1–5.4, 5.6 and 5.7 are satisfied. Then for any compact set* $\mathcal{K} \subset \mathring{\Delta}_d$*,* $\hat{\mathbb{A}}_n^{\mathrm{CFG}} \rightsquigarrow \hat{\mathbb{A}}^{\mathrm{CFG}}$ *as* $n \to \infty$ *in* $\mathcal{C}(\mathcal{K})$*, where for any* $\boldsymbol{w} \in \mathring{\Delta}_d$*,*

$$\hat{\mathbb{A}}^{\mathrm{CFG}}(\boldsymbol{w}) = A(\boldsymbol{w}) \int_0^1 \Big(\mathbb{C}[\psi_{\theta_0}\{-\boldsymbol{w}\log(u)\}] \\ + \sum_{j=1}^d \dot{C}_j[\psi_{\theta_0}\{-\boldsymbol{w}\log(u)\}] \dot{\psi}_{\theta_0}^\top\{-w_j \log(u)\}\Theta \Big) \frac{du}{u \log u}.$$

Theorem 5.2. *Under the assumptions of Theorem 5.1 with the additional assumption that* $s > 2$ *in case* ψ_{θ_0} *satisfies Condition 3.2 (a), and with Condition 5.4 replaced by Condition 5.5, one has that, for any compact set* $\mathcal{K} \subset \mathring{\Delta}_d$*,* $\hat{\mathbb{A}}_n^{\mathrm{P}} \rightsquigarrow \hat{\mathbb{A}}^{\mathrm{P}}$ *as* $n \to \infty$ *in* $\mathcal{C}(\mathcal{K})$*, where for any* $\boldsymbol{w} \in \mathring{\Delta}_d$*,*

$$\hat{\mathbb{A}}^{\mathrm{P}}(\boldsymbol{w}) = \frac{-A^2(\boldsymbol{w})}{\mathrm{E}(Z)} \int_0^1 \Big(\mathbb{C}[\psi_{\theta_0}\{-\boldsymbol{w}\log(u)\}] \\ + \sum_{j=1}^d \dot{C}_j[\psi_{\theta_0}\{-\boldsymbol{w}\log(u)\}] \dot{\psi}_{\theta_0}^\top\{-w_j \log(u)\}\Theta \Big) \frac{du}{u}.$$

With $\hat{\mu}$ and $\hat{\nu}$ as defined in Chapter 4, the end-point corrected versions of the CFG and Pickands-type estimators estimators are

$$\hat{A}_{n,c}^{\mathrm{P}}(\boldsymbol{w}) = n\hat{\mu} \Big/ \sum_{i=1}^n \hat{\xi}_{i,n}(\boldsymbol{w}), \quad \log \hat{A}_{n,c}^{\mathrm{CFG}}(\boldsymbol{w}) = \hat{\nu} - \frac{1}{n} \sum_{i=1}^n \log \hat{\xi}_{i,n}(\boldsymbol{w}). \tag{5.11}$$

By Lemma 4.1, the asymptotic behavior of the uncorrected and end-point corrected versions of the CFG and Pickands-type estimators is the same.

Corollary 5.1. *Theorems 5.1 and 5.2 also hold when* $\hat{\mathbb{A}}_n^{CFG}$ *and* $\hat{\mathbb{A}}_n^{\mathrm{P}}$ *are respectively replaced by* $\hat{\mathbb{A}}_{n,c}^{\mathrm{CFG}} = \sqrt{n}\,(\hat{A}_{n,c}^{\mathrm{CFG}} - A)$ *and* $\hat{\mathbb{A}}_{n,c}^{\mathrm{P}} = \sqrt{n}\,(\hat{A}_{n,c}^{\mathrm{P}} - A)$*.*

5.5 Proofs of Theorems 5.1 and 5.2

This section is devoted to the proof of Theorems 5.1 and 5.2. Consequences of the regularity conditions from Section 5.3 are first discussed in Section 5.5.1 and auxiliary results are gathered in Section 5.5.2. Theorems 5.1 and 5.2 are then proved in Sections 5.5.3 and 5.5.4, respectively.

5.5.1 Implications of the regularity conditions

First recall that it is assumed that θ_n is intrinsic, that is $\theta_n \in \mathcal{O}$ for all n. Expressions like ψ_{θ_n} and ϕ_{θ_n} are then well defined.

Under the conditions of either Theorem 5.1 or 5.2, Condition 5.2 implies that $(\bar{\mathbb{C}}_n, \Theta_n) \rightsquigarrow (\mathbb{C}, \Theta)$ in $\ell^\infty([0,1]^d) \times \mathbb{R}^p$ from Proposition 3.1 of Segers (2012). From Berghaus et al. (2017) it further follows that

$$(\bar{\mathbb{C}}_n/\tilde{g}_\omega, \Theta_n) \rightsquigarrow (\mathbb{C}/\tilde{g}_\omega, \Theta) \tag{5.12}$$

in $\ell^\infty([0,1]^d) \times \mathbb{R}^p$, where for any $\boldsymbol{u} \in [0,1]^d$, $\tilde{g}_\omega(\boldsymbol{u}) = g_\omega(\boldsymbol{u}) + \mathbb{1}\{g_\omega(\boldsymbol{u}) = 0\}$ for g_ω given by (4.10). Note that the requirement that Θ is Gaussian is actually not needed. In case Θ is centered but not Gaussian, the limiting process will be centered, but no longer Gaussian.

Next, Condition 5.4 implies that for each $j \in \{1, \ldots, p\}$ and $\theta \in \mathring{\mathcal{O}}$,

$$\int_0^{x_\Psi} \frac{|\dot{\psi}_{\theta,j}(t)|}{t} dt < \infty \quad \text{and} \quad \int_0^{x_\Psi} \frac{\|\dot{\psi}_\theta(t)\|_2^\omega}{t} dt < \infty \; ; \tag{5.13}$$

the latter holds because $\|\dot{\psi}_\theta(t)\|_2^\omega / h_\theta^\omega(t)$ is bounded on $[0, x_\Psi)$. Because h_θ is bounded, the same condition also implies that

$$\sup_{x \in [0, x_\Psi)} \|\dot{\psi}_\theta(x)\|_2 < \infty \tag{5.14}$$

and that

$$\lim_{\epsilon \downarrow 0} \sup_{\theta' \in \mathcal{O}, \|\theta' - \theta\|_2 \le \epsilon} \sup_{t \in [0, x_\Psi)} \|\dot{\psi}_{\theta'}(t) - \dot{\psi}_\theta(t)\|_2 = 0. \tag{5.15}$$

Moreover, given that for any $\omega \in (0,1)$ and any $a, b \ge 0$, $(a+b)^\omega \le a^\omega + b^\omega$, we have that $|a^\omega - b^\omega| \le |a - b|^\omega$. Hence, for any $t \in [0, x_\Psi)$ and $\theta, \theta' \in \mathring{\mathcal{O}}$,

$$\frac{\|\dot{\psi}_{\theta'}(t) - \dot{\psi}_\theta(t)\|_2^\omega}{h_\theta^\omega(t)} \ge \frac{|\|\dot{\psi}_{\theta'}(t)\|_2 - \|\dot{\psi}_\theta(t)\|_2|^\omega}{h_\theta^\omega(t)} \ge \frac{|\|\dot{\psi}_{\theta'}(t)\|_2^\omega - \|\dot{\psi}_\theta(t)\|_2^\omega|}{h_\theta^\omega(t)}$$

so that

$$\Upsilon_{\omega,\theta}(\epsilon) = \sup_{\theta' \in \mathcal{O}, \|\theta' - \theta\|_2 \le \epsilon} \sup_{t \in [0, x_\Psi)} \frac{|\|\dot{\psi}_{\theta'}(t)\|_2^\omega - \|\dot{\psi}_\theta(t)\|_2^\omega|}{h_\theta^\omega(t)} \to 0 \tag{5.16}$$

$\epsilon \to 0$. Similarly, Condition 5.5 implies (5.14), (5.15), and that (5.16), and that for each $j \in \{1, \ldots, d\}$ and $\theta \in \mathring{\mathcal{O}}$,

$$\int_0^{x_\Psi} |\dot{\psi}_{\theta,j}(t)| dt < \infty \quad \text{and} \quad \int_0^{x_\Psi} \|\dot{\psi}_\theta(t)\|_2^\omega dt < \infty. \tag{5.17}$$

5.5.2 Auxiliary results

As in Section 4.2.3, for an arbitrary $\boldsymbol{w} \in \Delta_d$, set $w_{(1)} = \min_{i=1,\ldots,d} w_i$ and $w_{(d)} = \max_{i=1,\ldots,d} w_i$. For any $k \in \mathbb{N}$, recall the set $B_{1/k} = \{\boldsymbol{w} \in \Delta_d : w_{(1)} \ge 1/k\}$.

Lemma 5.1. *Suppose that as $n \to \infty$, $\Theta_n = \sqrt{n}(\theta_n - \theta)$ converges in law to a nondegenerate limit Θ and that $\theta \in \mathring{\mathcal{O}}$. Further assume that Condition 5.3 holds and either Condition 5.4 or Condition 5.5 is satisfied. Then* *(i)* $\sqrt{n}(\psi_{\theta_n} - \psi_\theta) \rightsquigarrow \dot{\psi}_\theta^\top \Theta$ *as* $n \to \infty$ *in* $\mathcal{C}([0, x_\Psi))$.

(ii) $\sqrt{n}|\psi_{\theta_n} - \psi_\theta|/h_\theta \rightsquigarrow |\dot{\psi}_\theta^\top \Theta|/h_\theta$ *in* $\mathcal{C}([0, x_\Psi))$ *as* $n \to \infty$*, where* h_θ *is the weight function from Condition 5.4 and Condition 5.5, respectively, depending on which of these two conditions holds.*

(iii) If Condition 5.4 holds, then for any $0 \le a < b \le x_\Psi$*,* $\int_a^b \sqrt{n}|\psi_{\theta_n}(x) - \psi_\theta(x)|\frac{dx}{x} \rightsquigarrow \int_a^b |\dot{\psi}_\theta^\top(x)\Theta|\frac{dx}{x}$ *as* $n \to \infty$.

(iv) If Condition 5.5 holds, then for any $0 \le a < b \le x_\Psi$*,* $\int_a^b \sqrt{n}|\psi_{\theta_n}(x) - \psi_\theta(x)|dx \rightsquigarrow \int_a^b |\dot{\psi}_\theta^\top(x)\Theta|dx$ *as* $n \to \infty$.

Proof. (i). Because $\dot{\psi}_\theta$ is continuous by assumption and bounded in view of (5.14), $\dot{\psi}_\theta^\top \Theta_n \rightsquigarrow \dot{\psi}_\theta^\top \Theta$ as $n \to \infty$ in $\mathcal{C}([0, x_\Psi))$. Now let

$$Q_n = \sup_{x \in [0, x_\Psi)} |\sqrt{n}\{\psi_{\theta_n}(x) - \psi_\theta(x)\} - \dot{\psi}_\theta^\top(x)\Theta_n| \tag{5.18}$$

and choose an arbitrary $\varepsilon > 0$. Because (Θ_n) is tight, for any given $\delta > 0$ there exists $M_\delta > 0$ and $N_\delta > 0$ such that for all $n \ge N_\delta$, $\Pr(\|\Theta_n\|_2 > M_\delta) < \delta$. For any such n,

$$\Pr[Q_n > \varepsilon] < \Pr[Q_n > \varepsilon, \|\Theta_n\|_2 \le M_\delta] + \delta.$$

Suppose that n is large enough so that $\{\theta' \in \mathcal{O}, \|\theta' - \theta\|_2 \le n^{-1/2}M_\delta\} \subset \mathring{\mathcal{O}}$. Whenever $\|\theta_n - \theta\|_2 \le n^{-1/2}M_\delta$, an application of the Mean-Value Theorem implies that for every realization ϖ and $t \in [0, x_\Psi)$, $\psi_{\theta_n(\varpi)}(t) - \psi_\theta(t) = \dot{\psi}_{\Theta_n^*(t,\varpi)}^\top(t)(\theta_n(\varpi) - \theta)$, where $\Theta_n^*(t, \varpi) = \theta + \epsilon(t, \varpi)n^{-1/2}\Theta_n(\varpi)$ for some $\epsilon(t, \varpi) \in [0, 1]$. Hence,

$$\begin{aligned}
&\lim_{n\to\infty} \Pr[Q_n > \varepsilon, \|\Theta_n\|_2 \le M_\delta] \\
&\quad \le \lim_{n\to\infty} \Pr[\|\Theta_n\|_2 \sup_{x \in [0, x_\Psi)} \|\dot{\psi}_{\Theta_n^*(x)}(x) - \dot{\psi}_\theta(x)\|_2 > \varepsilon, \|\Theta_n\|_2 \le M_\delta] \\
&\quad \le \lim_{n\to\infty} \Pr[\sup_{\theta' \in \mathcal{O}, \|\theta' - \theta\| \le n^{-1/2}M_\delta} \sup_{x \in [0, x_\Psi)} \|\dot{\psi}_{\theta'}(x) - \dot{\psi}_\theta(x)\|_2 > \varepsilon/M_\delta] = 0,
\end{aligned}$$

where the last equality follows from (5.15). Given that δ can be chosen arbitrarily small, claim follows.

(ii). By the Continuous Mapping Theorem, $|\dot{\psi}_\theta^\top \Theta_n|/h_\theta \rightsquigarrow |\dot{\psi}_\theta^\top \Theta|/h_\theta$ in $\mathcal{C}([0, x_\Psi))$ as $n \to \infty$ given that for each $j \in \{1, \ldots, p\}$, $\dot{\psi}_{\theta,j}/h_\theta$ is bounded and continuous on $[0, x_\Psi)$ by Condition 5.4 or 5.5. It suffices to show that

$$V_n = \sup_{x \in [0, x_\psi)} \left| \frac{\sqrt{n}|\psi_{\theta_n}(x) - \psi_\theta(x)|}{h_\theta(x)} - \frac{|\dot{\psi}_\theta^\top(x)\Theta_n|}{h_\theta(x)} \right| \to 0$$

in probability as $n \to \infty$. As in the proof of (i), for any given $\delta > 0$ there exists $M_\delta > 0$ and $N_\delta > 0$ such that for all $n \ge N_\delta$, $\Pr(\|\Theta_n\|_2 > M_\delta) < \delta$. Suppose that $n \ge N_\delta$ is large enough so that $\{\theta' \in \mathcal{O}, \|\theta' - \theta\|_2 \le n^{-1/2}M_\delta\} \subset \mathring{\mathcal{O}}$. Whenever $\|\theta_n - \theta\|_2 \le n^{-1/2}M_\delta$, an application of the Mean-Value Theorem implies that for every realization ϖ and $t \in [0, x_\Psi)$, $\psi_{\theta_n(\varpi)}(t) - \psi_\theta(t) = \dot{\psi}^\top_{\Theta^*_n(t,\varpi)}(t)(\theta_n(\varpi) - \theta)$, where $\Theta^*_n(t,\varpi) = \theta + \epsilon(t,\varpi)n^{-1/2}\Theta_n(\varpi)$ for some $\epsilon(t,\varpi) \in [0,1]$. Hence,

$$\begin{aligned} V_n &= \sup_{x \in [0, x_\psi)} \left| \frac{|\dot{\psi}^\top_{\Theta^*_n(x)}(x)\Theta_n| - |\dot{\psi}^\top_\theta(x)\Theta_n|}{h_\theta(x)} \right| \\ &\leq \|\Theta_n\|_2 \sup_{x \in [0, x_\psi)} \frac{\|\dot{\psi}_{\Theta^*_n}(x) - \dot{\psi}_\theta(x)\|_2}{h_\theta(x)} \end{aligned}$$

For any such n and arbitrary $\varepsilon > 0$, $\Pr(V_n > \varepsilon)$ is at most

$$\delta + \Pr(V_n > \varepsilon, \|\Theta_n\|_2 \leq M_\delta) \leq \delta + \Pr\{M_\delta \Upsilon_\theta(M_\delta/\sqrt{n}) > \varepsilon\}$$

The second expression converges to 0 as $n \to \infty$ by Condition 5.4 or 5.5. Hence, $\lim_{n\to\infty} \Pr(V_n > \varepsilon) \leq \delta$. Since δ was arbitrary, the claim follows.

(iii) and (iv). This is a direct consequence of part (ii), the fact that either $\int_0^{x_\Psi} h_\theta(x)dx/x$ or $\int_0^{x_\Psi} h_\theta(x)dx$ is finite by assumption, as the case may be, and the Continuous Mapping Theorem. $\square$

Lemma 5.2. *Suppose that $n \to \infty$, $\Theta_n = \sqrt{n}(\theta_n - \theta)$ converges in law to a nondegenerate limit Θ and that $\theta \in \mathring{\mathcal{O}}$. Assume that Conditions 5.3,5.6 and 5.7 hold and that either Condition 5.4 or Condition 5.5 is satisfied. Then for any $c \in (0,1)$,*

(i) *As $n \to \infty$, $\phi_{\theta_n}(c/n) \to x_\Psi$ and $\phi_{\theta_n}(1 - c/n) \to 0$ in probability.*

(ii) *If Condition 3.2 holds for ψ_θ,*

$$\sqrt{n} \int_{\phi_{\theta_n}(c/n)}^{x_\Psi} \frac{\psi_\theta(x)}{x} dx$$

converges in probability to 0 as $n \to \infty$.

(iii) *If either Condition 3.2 (a) with $s > 2$, (b) or (c) holds for ψ_θ, then for any $c \in (0,1)$,*

$$\sqrt{n} \int_{\phi_{\theta_n}(c/n)}^{x_\Psi} \psi_\theta(x) dx$$

converges in probability to 0 as $n \to \infty$.

(iv) *If Condition 3.1 holds for ψ_θ, then for any $K \in \mathbb{N}$, $K \geq 2$,*

$$\sqrt{n} \int_{1/(K\phi_{\theta_n}(1-c/n))}^{\infty} \frac{1 - \psi_\theta(1/x)}{x^\ell} dx$$

converges in probability to 0 as $n \to \infty$, where $\ell \in \{1, 2\}$.

Proof. (i). It suffices to show that $\psi_\theta\{\phi_{\theta_n}(c/n)\} \to 0$ and $\psi_\theta\{\phi_{\theta_n}(1-c/n)\} \to 1$ in probability; the claim then follows from the Continuous Mapping Theorem. From the proof of Lemma 5.1 (i) we have that as $n \to \infty$, $Q_n \to 0$ in probability, where Q_n is as in (5.18). Also, recall from (5.14) that $M = \sup_{x\in[0,x_\Psi)} \|\dot{\psi}_\theta\|_2 < \infty$. Therefore, for arbitrary $\varepsilon > 0$,

$$\begin{aligned}
&\Pr[|\psi_\theta\{\phi_{\theta_n}(c/n)\}| > \varepsilon] \\
&\qquad \leq \Pr\{(c/n) > \varepsilon\} + \Pr[|\psi_{\theta_n}\{\phi_{\theta_n}(c/n)\} - \psi_\theta\{\phi_{\theta_n}(c/n)\}| > \varepsilon\}
\end{aligned}$$

and similarly

$$\begin{aligned}
&\Pr[|1-\psi_\theta\{\phi_{\theta_n}(1-c/n)\}| > \varepsilon] \\
&\qquad \leq \Pr\{(c/n) > \varepsilon\} + \Pr[|\psi_{\theta_n}\{\phi_{\theta_n}(1-c/n)\} - \psi_\theta\{\phi_{\theta_n}(1-c/n)\}| > \varepsilon\}
\end{aligned}$$

In both cases, the upper bound is at most

$$\Pr\{(c/n) > \varepsilon\} + \Pr(Q_n/\sqrt{n} > \varepsilon) + \Pr(M\|\theta_n - \theta\|_2 > \varepsilon)$$

which converges to 0 as $n \to \infty$.

(ii) and (iii). First, observe that

$$R_{n1} = \frac{1}{2}\Theta_n^\top \ddot{\psi}_\theta\{\phi_{\theta_n}(c/n)\}\Theta_n = o_P(1); \tag{5.19}$$

this follows readily from the Continuous Mapping Theorem, part (i) and Condition 5.6. Second, observe that

$$R_{n2} = \sqrt{n}\dot{\psi}_\theta^\top\{\phi_{\theta_n}(c/n)\}\Theta_n = o_P(1). \tag{5.20}$$

To show this, it suffices to prove that for any given $j \in \{1,\ldots,p\}$,

$$\sqrt{n}\dot{\psi}_{\theta,j}\{\phi_{\theta_n}(c/n)\} = o_P(1). \tag{5.21}$$

To this end, let $\ddot{\psi}_{\theta,j\bullet}$ denote the j-th row of the Hessian $\ddot{\psi}_\theta$. Because (Θ_n) is tight, for any given $\delta > 0$, there exists $M_\delta > 0$ and $N_\delta > 0$ such that for all $n \geq N_\delta$, $\Pr(\|\Theta_n\|_2 > M_\delta) < \delta$. Suppose that $n \geq N_\delta$ is large enough so that $\{\theta' \in \mathcal{O}, \|\theta'-\theta\|_2 \leq n^{-1/2}M_\delta\} \subset \mathring{\mathcal{O}}$. Whenever $\|\theta_n - \theta\|_2 \leq n^{-1/2}M_\delta$, the Mean-Value Theorem implies that for every realization ϖ and $t \in [0, x_\Psi)$, $\dot{\psi}_{\theta_n(\varpi),j}(t) - \dot{\psi}_{\theta,j}(t) = \ddot{\psi}^\top_{\Theta_n^*(t,\varpi),j\bullet}(t)(\theta_n(\varpi)-\theta)$, where $\Theta_n^*(t,\varpi) = \theta + \epsilon(t,\varpi)n^{-1/2}\Theta_n(\varpi)$ for some $\epsilon(t,\varpi) \in [0,1]$. Thus for any such n and arbitrary $\varepsilon > 0$,

$$\begin{aligned}
&\Pr(|\sqrt{n}[\dot{\psi}_{\theta_n,j}\{\phi_{\theta_n}(c/n)\} - \dot{\psi}_{\theta,j}\{\phi_{\theta_n}(c/n)\}] - \ddot{\psi}^\top_{\theta,j\bullet}\{\phi_{\theta_n}(c/n)\}\Theta_n| > \varepsilon) \\
&\qquad \leq \delta + \Pr(M_\delta \sup_{\theta'\in\mathcal{O}, \|\theta'-\theta\| < M_\delta/\sqrt{n}} \ \sup_{t\in[0,x_\Psi)} \|\ddot{\psi}_{\theta'}(t) - \ddot{\psi}_\theta(t)\|_{\mathrm{E}} > \varepsilon)
\end{aligned}$$

By Condition 5.6, the right-hand side converges to δ as $n \to \infty$. Because $\delta > 0$ was arbitrary,

$$|\sqrt{n}[\dot{\psi}_{\theta_n,j}\{\phi_{\theta_n}(c/n)\} \;-\; \dot{\psi}_{\theta,j}\{\phi_{\theta_n}(c/n)\}] \;-\; \ddot{\psi}^{\top}_{\theta,j\bullet}\{\phi_{\theta_n}(c/n)\}\Theta_n| \quad = \quad o_P(1). \quad (5.22)$$

Next, because for any $j, k \in \{1, \ldots, p\}$ $\ddot{\psi}_{\theta,jk}(x) \to 0$ as $x \to x_\Psi$ by Condition 5.6, part (i) implies that for arbitrary $j \in \{1, \ldots, p\}$,

$$\ddot{\psi}^{\top}_{\theta,j\bullet}\{\phi_{\theta_n}(c/n)\}\Theta_n = o_P(1). \quad (5.23)$$

Finally Condition 5.7 implies that for any $\delta > 0$ sufficiently small and arbitrary $\varepsilon > 0$,

$$\begin{aligned}&\Pr\{\sqrt{n}|\dot{\psi}_{\theta_n,j}\{\phi_{\theta_n}(c/n)\}| > \varepsilon\}\\ &\qquad\le \Pr(\|\theta_n - \theta\| > \delta) + \Pr\Big[\sqrt{c}\sup_{\theta':\|\theta'-\theta\|\le\delta}\sqrt{\frac{n}{c}}|\dot{\psi}_{\theta',j}\{\phi_{\theta'}(c/n)\}| > \varepsilon\Big]\end{aligned}$$

so that $\sqrt{n}\dot{\psi}_{\theta_n,j}\{\phi_{\theta_n}(c/n)\} = o_P(1)$. Combined with (5.23) and (5.22), we have that (5.21) holds for any $j \in \{1, \ldots, p\}$, and this in turns implies (5.20).

Next, observe that also

$$R_{n3} \;=\; \sup_{x\in[0,x_\Psi)} |n\{\psi_{\theta_n}(x) \;-\; \psi_\theta(x)\} \;-\; \sqrt{n}\dot{\psi}^{\top}_{\theta}(x)\Theta_n \;-\; \frac{1}{2}\Theta_n^{\top}\ddot{\psi}_\theta(x)\Theta_n| \;=\; o_P(1). \quad (5.24)$$

Indeed, by Taylor's Theorem with the mean-value remainder and the tightness of $\|\Theta_n\|_2$, for any $\varepsilon > 0$ and $\delta > 0$, and all $n \ge N_\delta$ large enough so that $\{\theta' \in \mathcal{O}, \|\theta' - \theta\|_2 \le n^{-1/2}M_\delta\} \subset \mathring{\mathcal{O}}$,

$$\Pr(R_{n3} > \varepsilon) \le \delta + \Pr\Big(\frac{M_\delta^2}{2}\sup_{\theta'\in\mathcal{O},\|\theta'-\theta\|\le M_\delta/\sqrt{n}}\;\sup_{t\in[0,x_\Psi)}\|\ddot{\psi}_{\theta'}(t) - \ddot{\psi}_\theta(t)\|_{\mathrm{E}} > \varepsilon\Big),$$

where $M_\delta, N_\delta > 0$ are such that for all $n \ge N_\delta$, $\Pr(\|\Theta_n\|_2 > M_\delta) < \delta$.

Putting all the pieces together, we have that

$$n|(c/n) - \psi_\theta\{\phi_{\theta_n}(c/n)\}| \le |R_{n1}| + |R_{n2}| + R_{n3} = o_P(1). \quad (5.25)$$

Whenever $n|(c/n) - \psi_\theta\{\phi_{\theta_n}(c/n)\}| \le \delta$ for some $\delta \in (0, \min\{c, 1-c\})$, the fact that ψ_θ is decreasing gives that

$$\phi_\theta\{(c+\delta)/n\} \le \phi_{\theta_n}(c/n) \le \phi_\theta\{(c-\delta)/n\}$$

Hence, for arbitrary $\varepsilon > 0$ and $\delta \in (0, \min\{c, 1-c\})$,

$$\begin{aligned}&\Pr\Big\{\sqrt{n}\int_{\phi_{\theta_n}(c/n)}^{x_\Psi}\frac{\psi_\theta(x)}{x}dx > \varepsilon\Big\} \le\\ &\qquad \Pr[n|(c/n) - \psi_\theta\{\phi_{\theta_n}(c/n)\}| > \delta] + \Pr\Big\{\sqrt{n}\int_{\phi_\theta((c+\delta)/n)}^{x_\Psi}\frac{\psi_\theta(x)}{x}dx > \varepsilon\Big\}.\end{aligned}$$

As $n \to \infty$, the first expression converges to 0 by (5.25), while the second converges to 0 by Lemma 4.5 (i). To establish part (iii), one can proceed exactly as above and conclude based on Lemma 4.5 (ii).

(iv). The proof is similar as that of part (ii). For,

$$n|(1-c/n) - \psi_\theta\{\phi_{\theta_n}(1-c/n)\}| \le |R^*_{n1}| + |R^*_{n2}| + R_{n3} = o_P(1), \tag{5.26}$$

where R_{n3} is as in (5.24),

$$R^*_{n1} = \frac{1}{2}\Theta_n^\top \ddot{\psi}_\theta\{\phi_{\theta_n}(1-c/n)\}\Theta_n = o_P(1)$$

from the Continuous Mapping Theorem, part (i) and Condition 5.7, and

$$R^*_{n2} = \sqrt{n}\dot{\psi}_\theta^\top\{\phi_{\theta_n}(1-c/n)\}\Theta_n = o_P(1)$$

using the same arguments as in the proof of part (ii) and Condition 5.7. Then for arbitrary $\varepsilon > 0$ and $\delta \in (0, \min(c, 1-c))$,

$$\begin{aligned}\Pr\Big\{\sqrt{n}\int_{1/\{K\phi_{\theta_n}(1-c/n)\}}^\infty \frac{1-\psi_\theta(1/x)}{x^\ell}dx > \varepsilon\Big\} \le \\ \Pr[n|(1-c/n) - \psi_\theta\{\phi_{\theta_n}(1-c/n)\}| > \delta] \\ + \Pr\Big\{\sqrt{n}\int_{1/[K\phi_\theta\{1-(c-\delta)/n\}]}^\infty \frac{1-\psi_\theta(1/x)}{x^\ell}dx > \varepsilon\Big\}.\end{aligned}$$

As $n \to \infty$, the first expression converges to 0 by (5.26), while the second converges to 0 by Lemma 4.5 (iii). □

Lemma 5.3. *Suppose that as $n \to \infty$, $\Theta_n = \sqrt{n}(\theta_n - \theta)$ converges in law to a nondegenerate limit Θ and that $\theta \in \mathring{\mathcal{O}}$. Further assume that Condition 5.3 holds and that either Condition 5.4 or 5.5 is satisfied. Then for any $K \in \mathbb{N}$, $K \ge 2$,*

(i) If Condition 5.4 holds,

$$\sup_{\boldsymbol{w}\in B_{1/K}} \int_0^{x_\Psi/w_{(d)}} |g_\omega\{\psi_{\theta_n}(\boldsymbol{w}x)\} - g_\omega\{\psi_\theta(\boldsymbol{w}x)\}|\frac{dx}{x} \to 0$$

in probability as $n \to \infty$.

(ii) If Condition 5.5 holds,

$$\sup_{\boldsymbol{w}\in B_{1/K}} \int_0^{x_\Psi/w_{(d)}} |g_\omega\{\psi_{\theta_n}(\boldsymbol{w}x)\} - g_\omega\{\psi_\theta(\boldsymbol{w}x)\}|dx \to 0$$

in probability as $n \to \infty$.

Proof. (i). Using the fact that for any $\omega \in (0, 1/2)$, the function t^ω on $[0, 1]$ is $C^{0,\omega}$ Hölder continuous and g_1 is Lipschitz continuous, there exist $\kappa_1, \kappa_2 > 0$ such that, for all $\boldsymbol{w} \in B_{1/K}$ and $x \in (0, x_\psi/w_{(d)})$,

$$|g_\omega\{\psi_{\theta_n}(\boldsymbol{w}x)\} - g_\omega\{\psi_\theta(\boldsymbol{w}x)\}| \leq \kappa_1\kappa_2^\omega \sum_{j=1}^d |\psi_{\theta_n}(w_jx) - \psi_\theta(w_jx)|^\omega,$$

Consequently,

$$\int_0^{x_\Psi/w_{(d)}} |g_\omega\{\psi_{\theta_n}(\boldsymbol{w}x)\} - g_\omega\{\psi_\theta(\boldsymbol{w}x)\}| \frac{dx}{x} \leq \kappa_1\kappa_2^\omega \sum_{j=1}^d \int_0^{x_\Psi/w_j} |\psi_{\theta_n}(w_jx) - \psi_\theta(w_jx)|^\omega \frac{dx}{x}.$$

By change of variable, the upper bound equals

$$\kappa_1\kappa_2^\omega d \int_0^{x_\Psi} |\psi_{\theta_n}(t) - \psi_\theta(t)|^\omega \frac{dt}{t}. \tag{5.27}$$

Whenever $\theta_n \in \mathring{\mathcal{O}}$, an application of the Mean-Value Theorem implies that for every realization ϖ and $t \in [0, x_\Psi)$, $|\psi_{\theta_n(\varpi)}(t) - \psi_\theta(t)| \leq \|\theta_n(\varpi) - \theta\|\|\dot{\psi}_{\Theta_n^*(t,\varpi)}(t)\|$, where $\Theta_n^*(t, \varpi) = \theta + \epsilon(t, \varpi)n^{-1/2}\Theta_n(\varpi)$ for some $\epsilon(t, \varpi) \in [0, 1]$. Consequently, (5.27) is bounded above by

$$\kappa_1\kappa_2^\omega d\|n^{-1/2}\Theta_n\|_2^\omega \int_0^{x_\Psi} \|\dot{\psi}_{\Theta_n^*(t)}(t)\|_2^\omega \frac{dt}{t},$$

which may be rewritten as

$$\kappa_1\kappa_2^\omega d\|n^{-1/2}\Theta_n\|_2^\omega \Big[\int_0^{x_\Psi} \|\dot{\psi}_\theta(t)\|_2^\omega \frac{dt}{t} + \int_0^{x_\Psi} \Big\{\|\dot{\psi}_{\Theta_n^*(t)}(t)\|_2^\omega - \|\dot{\psi}_\theta(t)\|_2^\omega\Big\} \frac{dt}{t}\Big].$$

Now fix an arbitrary $\epsilon > 0$ and $\delta > 0$ sufficiently small so that $\{\theta' \in \mathcal{O}, \|\theta' - \theta\|_2 \leq n^{-1/2}M_\delta\} \subset \mathring{\mathcal{O}}$. Then

$$\begin{aligned}\Pr&\left[\sup_{\boldsymbol{w}\in B_{1/K}} \int_0^{x_\Psi/w_{(d)}} |g_\omega\{\psi_{\theta_n}(\boldsymbol{w}x)\} - g_\omega\{\psi_\theta(\boldsymbol{w}x)\}| \frac{dx}{x} > \epsilon\right] \\ &\leq \Pr\left[\kappa_1\kappa_2^\omega d\delta^\omega \left\{\int_0^{x_\Psi} \|\dot{\psi}_\theta(t)\|_2^\omega \frac{dt}{t} + \Upsilon_{\omega,\theta}(\delta) \int_0^{x_\Psi} h_\theta^\omega(t) \frac{dt}{t}\right\} > \epsilon\right] \\ &\quad + \Pr(\|n^{-1/2}\Theta_n\|_2 > \delta),\end{aligned}$$

where $\Upsilon_{\omega,\theta}$ is as in (5.16). Since $\Pr(\|n^{-1/2}\Theta_n\|_2 > \delta) \to 0$ as $n \to \infty$, one has, for any $\delta > 0$,

$$\lim_{n\to\infty} \Pr\left[\sup_{\boldsymbol{w}\in B_{1/K}} \int_0^{x_\Psi/w_{(d)}} |g_\omega\{\psi_{\theta_n}(\boldsymbol{w}x)\} - g_\omega\{\psi_\theta(\boldsymbol{w}x)\}| \frac{dx}{x} > \epsilon\right] \leq \Pr\left[\kappa_1\kappa_2^\omega d\delta^\omega \left\{\int_0^{x_\Psi} \|\dot{\psi}_\theta(t)\|_2^\omega \frac{dt}{t} + \Upsilon_{\omega,\theta}(\delta) \int_0^{x_\Psi} h_\theta^\omega(t) \frac{dt}{t}\right\} > \epsilon\right].$$

The right-hand side converges to 0 as $\delta \to 0$. Indeed, (5.16) implies that

$$\lim_{\delta \downarrow 0} \left[\kappa_1 \kappa_2^\omega d \delta^\omega \left\{ \int_0^{x_\Psi} \|\dot{\psi}_\theta(t)\|_2^\omega \frac{dt}{t} + \Upsilon_{\omega,\theta}(\delta) \int_0^{x_\Psi} h_\theta^\omega(t) \frac{dt}{t} \right\} \right] = 0$$

given that both integrals are finite by Condition 5.4 and (5.13).

(ii). The proof is completely analogous to the proof of (i). Using the same arguments, there exist constants $\kappa_1, \kappa_2 > 0$ such that for $\boldsymbol{w} \in B_{1/K}$,

$$\int_0^{x_\Psi / w_{(d)}} |g_\omega\{\psi_{\theta_n}(\boldsymbol{w}x)\} - g_\omega\{\psi_\theta(\boldsymbol{w}x)\}| dx \leq \kappa_1 \kappa_2^\omega d K \int_0^{x_\Psi} |\psi_{\theta_n}(t) - \psi_\theta(t)|^\omega dt.$$

One can proceed as above using Condition 5.5 and (5.17). □

5.5.3 Proof of Theorem 5.1

Let $\mathcal{K}$ be a compact subset of $\mathring{\Delta}_d$. For an arbitrary $\boldsymbol{w} \in \Delta_d$, set $w_{(1)} = \min_{i=1,\ldots,d} w_i$ and $w_{(d)} = \max_{i=1,\ldots,d} w_i$. Define, for any $k \in \mathbb{N}$, the set $B_{1/k} = \{\boldsymbol{w} \in \Delta_d : w_{(1)} \geq 1/k\}$. Since $\mathcal{K}$ is compact, there exists an integer $K > 1$ such that $\mathcal{K} \subset B_{1/K} \subset \mathring{\Delta}_d$.

To simplify notation, we denote the true parameter value by θ instead of θ_0 henceforth and set $\Theta_n = \sqrt{n}(\theta_n - \theta)$.

As in Section 4.1, introduce the process $\hat{\mathbb{B}}_n^{\mathrm{CFG}}$ given, for all $\boldsymbol{w} \in \Delta_d$, by

$$\hat{\mathbb{B}}_n^{\mathrm{CFG}}(\boldsymbol{w}) = \sqrt{n} \left\{ \log \hat{A}_n^{\mathrm{CFG}}(\boldsymbol{w}) - \log A(\boldsymbol{w}) \right\}.$$

Proceeding as in the proof of Lemma 4.2, $\hat{\mathbb{B}}_n^{\mathrm{CFG}}$ may be rewritten as

$$\begin{aligned} \hat{\mathbb{B}}_n^{\mathrm{CFG}}(\boldsymbol{w}) &= -\int_0^\infty \sqrt{n} \left[\hat{C}_n\{\psi_{\theta_n}(\boldsymbol{w}x)\} - C\{\psi_\theta(\boldsymbol{w}x)\} \right] \frac{dx}{x} \\ &= -\int_0^{x_\Psi / w_{(d)}} \sqrt{n} \left[\hat{C}_n\{\psi_{\theta_n}(\boldsymbol{w}x)\} - C\{\psi_\theta(\boldsymbol{w}x)\} \right] \frac{dx}{x} \end{aligned}$$

where the second equality follows because $\hat{C}_n\{\psi_{\theta_n}(\boldsymbol{w}x)\} = C\{\psi_\theta(\boldsymbol{w}x)\} = 0$ whenever $x > x_\Psi / w_{(d)}$ if Condition 5.3 holds. Next, write $\hat{\mathbb{B}}_n^{\mathrm{CFG}} = \hat{\mathbb{B}}_{n1}^{\mathrm{CFG}} + \hat{\mathbb{B}}_{n2}^{\mathrm{CFG}}$, where for all $\boldsymbol{w} \in \Delta_d$,

$$\hat{\mathbb{B}}_{n1}^{\mathrm{CFG}}(\boldsymbol{w}) = -\int_0^{x_\Psi / w_{(d)}} \hat{\mathbb{C}}_n\{\psi_{\theta_n}(\boldsymbol{w}x)\} \frac{dx}{x}$$

and

$$\hat{\mathbb{B}}_{n2}^{\mathrm{CFG}}(\boldsymbol{w}) = -\int_0^{x_\Psi / w_{(d)}} \sqrt{n} \left[C\{\psi_{\theta_n}(\boldsymbol{w}x)\} - C\{\psi_\theta(\boldsymbol{w}x)\} \right] \frac{dx}{x}. \tag{5.28}$$

For reasons that will become apparent in the proof of Lemma 5.6 below, it is important to first establish the asymptotic behavior of the drift $\hat{\mathbb{B}}_{n2}^{\mathrm{CFG}}$. To this end, let $\breve{\mathbb{B}}_{n2}^{\mathrm{CFG}}$ be the process given for all $\boldsymbol{w} \in \mathring{\Delta}_d$ by $\breve{\mathbb{B}}_{n2}^{\mathrm{CFG}}(\boldsymbol{w}) = a^\top(\boldsymbol{w}) \Theta_n$ where $a(\boldsymbol{w}) = (a_1(\boldsymbol{w}), \ldots, a_p(\boldsymbol{w}))^\top$ with

$$a_k(\boldsymbol{w}) = -\sum_{j=1}^d \int_0^{x_\Psi / w_j} \dot{C}_j\{\psi_\theta(\boldsymbol{w}x)\} \dot{\psi}_{\theta,k}(w_j x) \frac{dx}{x}. \tag{5.29}$$

The following lemma establishes that $|a_k(\boldsymbol{w})| < \infty$ for any $k \in \{1, \ldots, p\}$ and $\boldsymbol{w} \in \mathring{\Delta}_d$, and specifies the weak limit of $\check{\mathbb{B}}_{n2}^{\mathrm{CFG}}$.

Lemma 5.4. *As $n \to \infty$, $\check{\mathbb{B}}_{n2}^{\mathrm{CFG}} \rightsquigarrow \mathbb{B}_2^{\mathrm{CFG}}$ in $\mathcal{C}(B_{1/K})$, where for all $\boldsymbol{w} \in \mathring{\Delta}_d$, $\mathbb{B}_2^{\mathrm{CFG}}(\boldsymbol{w}) = a^\top(\boldsymbol{w})\Theta$.*

Proof. First, note that for any $k \in \{1, \ldots, p\}$ and $\boldsymbol{w} \in \mathring{\Delta}_d$, $|a_k(\boldsymbol{w})| < \infty$. Indeed, since $0 \le \dot{C}_j \le 1$ for all $j \in \{1, \ldots, d\}$,

$$\begin{aligned} |a_k(\boldsymbol{w})| &\le \sum_{j=1}^d \int_0^{x_\Psi/w_j} \dot{C}_j\{\psi_\theta(\boldsymbol{w}x)\}|\dot{\psi}_{\theta,k}(w_j x)|\frac{dx}{x} \\ &\le \sum_{j=1}^d \int_0^{x_\Psi/w_j} |\dot{\psi}_{\theta,k}(w_j x)|\frac{dx}{x} = d\int_0^{x_\Psi} |\dot{\psi}_{\theta,k}(t)|\frac{dt}{t}. \end{aligned}$$

The last expression is finite by Condition 5.4. The next step is to show that a is uniformly continuous on $B_{1/K}$, viz.

$$\lim_{\delta\downarrow 0} \sup_{\boldsymbol{w},\boldsymbol{w}' \in B_{1/K}, \|\boldsymbol{w}-\boldsymbol{w}'\|_2<\delta} \|a(\boldsymbol{w}) - a(\boldsymbol{w}')\|_2 = 0. \tag{5.30}$$

To show that (5.30) holds, define, for all $j \in \{1, \ldots, d\}$ and $k \in \{1, \ldots, p\}$,

$$b_{j,k}(\boldsymbol{w}) = \int_0^{x_\Psi/w_j} \dot{C}_j\{\psi_\theta(\boldsymbol{w}x)\}\dot{\psi}_{\theta,k}(w_j x)\frac{dx}{x}.$$

Then (5.30) follows if for all $j \in \{1, \ldots, d\}$ and $k \in \{1, \ldots, p\}$,

$$\lim_{\delta\downarrow 0} \sup_{\boldsymbol{w},\boldsymbol{w}' \in B_{1/K}, \|\boldsymbol{w}-\boldsymbol{w}'\|_2<\delta} |b_{j,k}(\boldsymbol{w}) - b_{j,k}(\boldsymbol{w}')| = 0.$$

Pick an arbitrary $j \in \{1, \ldots, d\}$, $k \in \{1, \ldots, p\}$. Then for any $\boldsymbol{w}, \boldsymbol{w}' \in B_{1/K}$,

$$|b_{j,k}(\boldsymbol{w}) - b_{j,k}(\boldsymbol{w}')| \le \int_0^{x_\Psi} |\dot{C}_j\{\psi_\theta(\boldsymbol{w}t/w_j)\} - \dot{C}_j\{\psi_\theta(\boldsymbol{w}'t/w_j')\}||\dot{\psi}_{\theta,k}(t)|\frac{dt}{t}$$

by the change of variable. Now pick an arbitrary $\eta, \mu \in (0, 1)$ and note that because ψ_θ is uniformly continuous, there exists $\lambda > 0$ such that for all $|x - x'| < \lambda$, $|\psi_\theta(x) - \psi_\theta(x')| < \mu$. Also note that if $\|\boldsymbol{w} - \boldsymbol{w}'\|_2 < \delta$, $|(w_k t/w_j) - (w_k' t/w_j')| \le 2K^2 t\delta$. Because $2K^2\phi_\theta(\eta)\delta < \lambda$ for all δ sufficiently close to 0 and because $0 \le \dot{C}_j \le 1$,

$$\begin{aligned} \lim_{\delta\downarrow 0} \sup_{\substack{\boldsymbol{w},\boldsymbol{w}' \in B_{1/K},\\ \|\boldsymbol{w}-\boldsymbol{w}'\|_2<\delta}} |b_{j,k}(\boldsymbol{w}) - b_{j,k}(\boldsymbol{w}')| &\le \sup_{\substack{\boldsymbol{u},\boldsymbol{u}' \in A_{\eta,j}\\ \|\boldsymbol{u}-\boldsymbol{u}'\|_2<\mu}} |\dot{C}_j(\boldsymbol{u}) - \dot{C}_j(\boldsymbol{u}')| \int_{\phi_\theta(1-\eta)}^{\phi_\theta(\eta)} |\dot{\psi}_{\theta,k}(t)|\frac{dt}{t} \\ &\quad + 2\Big\{\int_0^{\phi_\theta(1-\eta)} |\dot{\psi}_{\theta,k}(t)|\frac{dt}{t} + \int_{\phi_\theta(\eta)}^{x_\Psi} |\dot{\psi}_{\theta,k}(t)|\frac{dt}{t}\Big\}, \end{aligned}$$

where $A_{\eta,j} = \{\boldsymbol{u} \in [0,1]^d : u_j \in [\eta, 1-\eta]\}$. Because $\dot{C}_j$ is uniformly continuous on the set $A_{\eta,j}$ by Proposition 3.2, the first expression on the right-hand side tends to 0 as $\mu \to 0$. Because $\int_0^{x_\Psi}\{|\dot{\psi}_{\theta,k}(t)|\}/t dt$ is finite by Condition 5.4, the second expression tends to 0 as $\eta \to 0$. $\square$

The next step is to establish, through the following lemma, that the limiting behavior of $\hat{\mathbb{B}}_{n2}^{\mathrm{CFG}}$ is the same as that of $\check{\mathbb{B}}_{n2}^{\mathrm{CFG}}$.

Lemma 5.5. *As $n \to \infty$, $\sup_{\boldsymbol{w} \in B_{1/K}} |\hat{\mathbb{B}}_{n2}^{\mathrm{CFG}}(\boldsymbol{w}) - \check{\mathbb{B}}_{n2}^{\mathrm{CFG}}(\boldsymbol{w})|$ converges in probability to* 0.

Proof. Let $\tilde{\mathbb{B}}_{n2}^{\mathrm{CFG}}$ be given, for all $\boldsymbol{w} \in \mathring{\Delta}_d$, by

$$\tilde{\mathbb{B}}_{n2}^{\mathrm{CFG}}(\boldsymbol{w}) = -\sum_{j=1}^{d} \int_0^{x_\Psi/w_j} \sqrt{n}\{\psi_{\theta_n}(w_j x) - \psi_\theta(w_j x)\} \dot{C}_j\{\psi_\theta(\boldsymbol{w}x)\} \frac{dx}{x}.$$

We will first show that

$$\sup_{\boldsymbol{w} \in B_{1/K}} |\tilde{\mathbb{B}}_{n2}^{\mathrm{CFG}}(\boldsymbol{w}) - \hat{\mathbb{B}}_{n2}^{\mathrm{CFG}}(\boldsymbol{w})| = o_P(1). \tag{5.31}$$

Using the Mean-Value Theorem, write

$$\begin{aligned}
\hat{\mathbb{B}}_{n2}^{\mathrm{CFG}}(\boldsymbol{w}) &= -\int_0^{x_\Psi/w_{(d)}} \sqrt{n}\left[C\{\psi_{\theta_n}(\boldsymbol{w}x)\} - C\{\psi_\theta(\boldsymbol{w}x)\}\right] \frac{dx}{x} \\
&= -\sum_{j=1}^{d} \int_0^{x_\Psi/w_j} \sqrt{n}\{\psi_{\theta_n}(w_j x) - \psi_\theta(w_j x)\} \dot{C}_j(\boldsymbol{u}_{\boldsymbol{w}x}) \frac{dx}{x},
\end{aligned}$$

where for every $\boldsymbol{w}x$ and realization ϖ, $\boldsymbol{u}_{\boldsymbol{w}x}(\varpi) = \epsilon(\boldsymbol{w}x, \varpi)\psi_{\theta_n(\varpi)}(\boldsymbol{w}x) + \{1 - \epsilon(\boldsymbol{w}x, \varpi)\}$ $\psi_\theta(\boldsymbol{w}x)$ for some $\epsilon(\boldsymbol{w}x, \varpi) \in [0, 1]$. It thus suffices to show that for all $j \in \{1, \ldots, d\}$,

$$V_n = \sup_{\boldsymbol{w} \in B_{1/K}} \int_0^{x_\Psi/w_j} \sqrt{n}|\psi_{\theta_n}(w_j x) - \psi_\theta(w_j x)||\dot{C}_j\{\psi_\theta(\boldsymbol{w}x)\} - \dot{C}_j(\boldsymbol{u}_{\boldsymbol{w}x})| \frac{dx}{x}$$

converges in probability to 0 as $n \to \infty$. To accomplish this, fix an arbitrary $j \in \{1, \ldots, d\}$ and let

$$T_n = \sup_{x \in [0, x_\Psi)} |\sqrt{n}\{\psi_{\theta_n}(x) - \psi_\theta(x)\}| \,. \tag{5.32}$$

From Lemma 5.1 (i), it follows that the sequence (T_n) is tight. For any $\delta > 0$ there exists $M_\delta > 0$ and $N_\delta > 0$ such that for all $n \geq N_\delta$, $\Pr(T_n > M_\delta) < \delta$. Pick an an arbitrary $\varepsilon > 0$, $\eta \in (0, 1)$ and let $n \geq N_\delta$ be such that $M_\delta/\sqrt{n} < \eta/2$. Then $\Pr(V_n > \varepsilon) \leq \delta + \Pr(V_n > \varepsilon, T_n \leq M_\delta)$ and $\Pr(V_n > \varepsilon, T_n \leq M_\delta)$ may be bounded above by $\Pr(V_{n1} > \varepsilon/2) + \Pr(V_{n2} > \varepsilon/2)$, where

$$\begin{aligned}
V_{n1} = 2 \sup_{\boldsymbol{w} \in B_{1/K}} &\Big\{ \int_0^{\phi_\theta(1-\eta)/w_j} \sqrt{n}|\psi_{\theta_n}(w_j x) - \psi_\theta(w_j x)| \frac{dx}{x} \\
&+ \int_{\phi_\theta(\eta)/w_j}^{x_\Psi/w_j} \sqrt{n}|\psi_{\theta_n}(w_j x) - \psi_\theta(w_j x)| \frac{dx}{x} \Big\} \\
= 2 &\Big\{ \int_0^{\phi_\theta(1-\eta)} \sqrt{n}|\psi_{\theta_n}(t) - \psi_\theta(t)| \frac{dt}{t} + \int_{\phi_\theta(\eta)}^{x_\Psi} \sqrt{n}|\psi_{\theta_n}(t) - \psi_\theta(t)| \frac{dt}{t} \Big\}
\end{aligned}$$

and

$$V_{n2} = \sup_{\substack{\boldsymbol{u},\boldsymbol{u}' \in A_{\eta/2,j} \\ \|\boldsymbol{u}-\boldsymbol{u}'\| < M_\delta/\sqrt{n}}} |\dot{C}_j(\boldsymbol{u}) - \dot{C}_j(\boldsymbol{u}')| \int_{\phi_\theta(1-\eta)}^{\phi_\theta(\eta)} \sqrt{n}|\psi_{\theta_n}(t) - \psi_\theta(t)| \frac{dt}{t},$$

where $A_{\eta/2,j} = \{\boldsymbol{u} \in [0,1]^d : u_j \in [\eta/2, 1-\eta/2]\}$. Because $\dot{C}_j$ is uniformly continuous on $A_{\eta/2,j}$ and

$$\int_{\phi_\theta(1-\eta)}^{\phi_\theta(\eta)} \sqrt{n}|\psi_{\theta_n}(x) - \psi_\theta(x)| \frac{dx}{x} \rightsquigarrow \int_{\phi_\theta(1-\eta)}^{\phi_\theta(\eta)} |\dot{\psi}_\theta^\top(x)\Theta| \frac{dx}{x}$$

as $n \to \infty$ by Lemma 5.1 (iii), $V_{n2} \to 0$ in probability as $n \to \infty$. The same lemma, again part (iii), also implies that as $n \to \infty$,

$$V_{n1} \rightsquigarrow 2\Big\{ \int_0^{\phi_\theta(1-\eta)} |\dot{\psi}_\theta^\top(x)\Theta| \frac{dx}{x} + \int_{\phi_\theta(\eta)}^{x_\Psi} |\dot{\psi}_\theta^\top(x)\Theta| \frac{dx}{x} \Big\}.$$

The limit is non-negative and bounded above by

$$2 \sup_{t \in [0, x_\Psi)} \frac{|\dot{\psi}_\theta^\top(t)\Theta|}{h_\theta(t)} \Big\{ \int_0^{\phi_\theta(1-\eta)} h_\theta(x) \frac{dx}{x} + \int_{\phi_\theta(\eta)}^{x_\Psi} h_\theta(x) \frac{dx}{x} \Big\} .$$

By the Portmanteau Lemma, the lim sup of $\Pr(V_{n1} > \varepsilon/2)$ is at most

$$\Pr\Big[2 \sup_{t \in [0, x_\Psi)} \frac{|\dot{\psi}_\theta^\top(t)\Theta|}{h_\theta(t)} \Big\{ \int_0^{\phi_\theta(1-\eta)} h_\theta(x) \frac{dx}{x} + \int_{\phi_\theta(\eta)}^{x_\Psi} h_\theta(x) \frac{dx}{x} \Big\} \ge \varepsilon/2 \Big] .$$

This probability can be made arbitrarily small given that

$$\lim_{\eta \to 0} \Big\{ \int_0^{\phi_\theta(1-\eta)} h_\theta(x) \frac{dx}{x} + \int_{\phi_\theta(\eta)}^{x_\Psi} h_\theta(x) \frac{dx}{x} \Big\} = 0 .$$

Since δ was arbitrary, $\Pr(V_n > \varepsilon) \to 0$ as $n \to \infty$, and (5.31) holds.

Next, we establish that

$$\sup_{\boldsymbol{w} \in B_{1/K}} |\tilde{\mathbb{B}}_{n2}^{\mathrm{CFG}}(\boldsymbol{w}) - \breve{\mathbb{B}}_{n2}^{\mathrm{CFG}}(\boldsymbol{w})| = o_P(1) . \tag{5.33}$$

To this end, it suffices to show that for each $j \in \{1, \ldots, d\}$,

$$\sup_{\boldsymbol{w} \in B_{1/K}} \int_0^{x_\Psi/w_j} |\dot{C}_j\{\psi_\theta(\boldsymbol{w}x)\}| |\dot{\psi}_\theta^\top(w_j x)\Theta_n - \sqrt{n}\{\psi_{\theta_n}(w_j x) - \psi_\theta(w_j x)\}| \frac{dx}{x}$$

converges to 0 in probability as $n \to \infty$. Using the fact that $0 \le \dot{C}_j \le 1$ and making a change of variable, this expression is bounded above by

$$W_n = \int_0^{x_\Psi} |\dot{\psi}_\theta^\top(t)\Theta_n - \sqrt{n}\{\psi_{\theta_n}(t) - \psi_\theta(t)\}| \frac{dt}{t} .$$

We can now proceed similarly as in the proof of (ii) of Lemma 5.1. Because (Θ_n) is tight, for any given $\delta > 0$, there exists $M_\delta > 0$ and $N_\delta > 0$ such that for all $n \ge N_\delta$,

$\Pr(\|\Theta_n\|_2 > M_\delta) < \delta$. Suppose that $n \ge N_\delta$ is large enough so that $\{\theta' \in \mathcal{O}, \|\theta' - \theta\|_2 \le n^{-1/2}M_\delta\} \subset \mathring{\mathcal{O}}$. Whenever $\|\theta_n - \theta\|_2 \le n^{-1/2}M_\delta$,

$$W_n \le \|\Theta_n\|_2 \sup_{x\in[0,x_\psi)} \frac{\|\dot\psi_{\Theta_n^*(x)}(x) - \dot\psi_\theta(x)\|_2}{h_\theta(x)} \int_0^{x_\Psi} h_\theta(t)\frac{dt}{t},$$

where for any realization ϖ, $\Theta_n^*(x,\varpi) = \theta + \epsilon(x,\varpi)n^{-1/2}\Theta_n(\varpi)$ for some $\epsilon(x,\varpi) \in [0,1]$. For any such n and arbitrary $\varepsilon > 0$, $\Pr(W_n > \varepsilon)$ is at most

$$\delta + \Pr(W_n > \varepsilon, \|\Theta_n\|_2 \le M_\delta) \le \delta + \Pr\Big\{M_\delta \Upsilon_\theta(M_\delta/\sqrt{n}) \int_0^{x_\Psi} h_\theta(t)\frac{dt}{t} > \varepsilon\Big\}$$

Clearly, the second expression converges to 0 as $n \to \infty$ by Condition 5.4. Hence, $\lim_{n\to\infty} \Pr(W_n > \varepsilon) \le \delta$. Since δ was arbitrary, (5.33) follows. □

Combining Lemmas 5.4 and 5.5, we thus have that

$$\hat{\mathbb{B}}_{n2}^{\mathrm{CFG}} \rightsquigarrow \mathbb{B}_2^{\mathrm{CFG}} \tag{5.34}$$

as $n \to \infty$ in $\mathcal{C}(B_{1/K})$, where for all $\boldsymbol{w} \in \mathring{\Delta}_d$, $\mathbb{B}_2^{\mathrm{CFG}}(\boldsymbol{w}) = a^\top(\boldsymbol{w})\Theta$. Next, let $\bar{\mathbb{C}}_n$ be as in Theorem 2.12 in Chapter 2 and define for all $\boldsymbol{w} \in \Delta_d$,

$$\bar{\mathbb{B}}_{n1}^{\mathrm{CFG}}(\boldsymbol{w}) = -\int_0^{x_\Psi/w_{(d)}} \bar{\mathbb{C}}_n\{\psi_{\theta_n}(\boldsymbol{w}x)\}\frac{dx}{x}. \tag{5.35}$$

The following lemma is the analogue of Lemma 4.6.

Lemma 5.6. *As $n \to \infty$,*

$$\sup_{\boldsymbol{w}\in B_{1/K}} \int_0^{x_\Psi/w_{(d)}} |\hat{\mathbb{C}}_n\{\psi_{\theta_n}(\boldsymbol{w}x)\} - \bar{\mathbb{C}}_n\{\psi_{\theta_n}(\boldsymbol{w}x)\}|\frac{dx}{x}$$

converges in probability to 0.

Proof. First, pick an arbitrary $c \in (0, 1/K)$ and define

$$a_n = \phi_{\theta_n}\left(1 - \frac{c}{n}\right), \quad b_n = \phi_{\theta_n}\left(\frac{c}{n}\right).$$

Let $N_K \in \mathbb{N}$ be such that for any $n \ge N_K$, $c < n/\{K(n+1)\}$. Throughout the proof, assume that $n \ge N_K$. Then $c < \frac{n}{n+1}$ and, by Lemma 4.3,

$$\psi_{\theta_n}\left\{K\phi_{\theta_n}\left(1 - \frac{c}{n}\right)\right\} > \frac{n}{n+1}. \tag{5.36}$$

As in the proof of Lemma 4.6, use the triangle inequality to write

$$\int_0^{x_\Psi/w_{(d)}} |\hat{\mathbb{C}}_n\{\psi_{\theta_n}(\boldsymbol{w}x)\} - \bar{\mathbb{C}}_n\{\psi_{\theta_n}(\boldsymbol{w}x)\}|\frac{dx}{x} \le I_1(\boldsymbol{w}) + I_2(\boldsymbol{w}) + I_3(\boldsymbol{w}) + I_4(\boldsymbol{w}) + I_5(\boldsymbol{w}),$$

with

$$I_1(\boldsymbol{w}) = \int_{a_n/w_{(1)}}^{b_n/w(d)} \left|\hat{\mathbb{C}}_n\{\psi_{\theta_n}(\boldsymbol{w}x) - \bar{\mathbb{C}}_n\{\psi_{\theta_n}(\boldsymbol{w}x)\right| \frac{dx}{x},$$

$$I_2(\boldsymbol{w}) = \int_0^{a_n/w_{(1)}} \left|\hat{\mathbb{C}}_n\{\psi_{\theta_n}(\boldsymbol{w}x)\right| \frac{dx}{x},\ I_3(\boldsymbol{w}) = \int_{b_n/w(d)}^{x_\Psi/w_{(d)}} \left|\hat{\mathbb{C}}_n\{\psi_{\theta_n}(\boldsymbol{w}x)\right| \frac{dx}{x},$$

$$I_4(\boldsymbol{w}) = \int_0^{a_n/w_{(1)}} \left|\bar{\mathbb{C}}_n\{\psi_{\theta_n}(\boldsymbol{w}x)\right| \frac{dx}{x},\ I_5(\boldsymbol{w}) = \int_{b_n/w(d)}^{x_\Psi/w_{(d)}} \left|\bar{\mathbb{C}}_n\{\psi_{\theta_n}(\boldsymbol{w}x)\right| \frac{dx}{x}\ .$$

Each integral will be treated separately, showing that for all $p \in \{1, \ldots, 5\}$, $\sup_{\boldsymbol{w}\in B_{1/K}} I_p(\boldsymbol{w}) \to 0$ in probability as $n \to \infty$.

Treatment of I_1. Fix $\boldsymbol{w} \in B_{1/K}$ and let g_ω be the weight function given by (4.10) for any $\omega \in (0, 1/2)$. Since $a_n/w_{(1)} < x < b_n/w_{(d)}$, $c/n < \psi_{\theta_n}(w_j x) < 1-c/n$ for all $j \in \{1, \ldots, d\}$. Thus with S_n as in (4.11),

$$\begin{aligned}
I_1(\boldsymbol{w}) &\le S_n \int_0^{x_\Psi/w_{(d)}} \frac{g_\omega\{\psi_{\theta_n}(\boldsymbol{w}x)\}}{x} dx \\
&\le S_n \left[\int_0^{x_\Psi/w_{(d)}} \frac{g_\omega\{\psi_\theta(\boldsymbol{w}x)\}}{x} dx \right. \\
&\qquad \left. + \sup_{\boldsymbol{w}\in B_{1/K}} \int_0^{x_\Psi/w_{(d)}} |g_\omega\{\psi_{\theta_n}(\boldsymbol{w}x)\} - g_\omega\{\psi_\theta(\boldsymbol{w}x)\}| \frac{dx}{x}\right] .
\end{aligned}$$

By the first part of Theorem 2.12, S_n converges to 0 in probability as $n \to \infty$, while Lemma 5.3 implies that the expression in the square brackets converges in probability to

$$\int_0^{x_\Psi/w_{(d)}} \frac{g_\omega\{\psi_\theta(\boldsymbol{w}x)\}}{x} dx,$$

which was shown to be finite while discussing I_1 in the proof of Lemma 4.6.

Treatment of I_2. Fixing an arbitrary $\boldsymbol{w} \in B_{1/K}$, for any $x \in (0, a_n/w_{(1)})$ and $j \in \{1, \ldots, d\}$, $w_j x \le (w_j/w_{(1)})\phi_{\theta_n}(1-c/n) \le K\phi_{\theta_n}(1-c/n)$. Together with (5.36), this implies that $\psi_{\theta_n}(w_j x) \ge \psi_{\theta_n}\{K\phi_{\theta_n}(1-c/n)\} > n/(n+1)$. Therefore, for any $x \in (0, a_n/w_{(1)})$, $\hat{C}_n\{\psi_{\theta_n}(\boldsymbol{w}x)\} = 1$ and $I_2(\boldsymbol{w}) = I_{21}(\boldsymbol{w}) + I_{22}(\boldsymbol{w})$, where

$$\begin{aligned}
I_{21}(\boldsymbol{w}) &= \sqrt{n} \int_0^{a_n/w_{(1)}} [1 - C\{\psi_\theta(\boldsymbol{w}x)\}] \frac{dx}{x}, \\
I_{22}(\boldsymbol{w}) &= \sqrt{n} \int_0^{a_n/w_{(1)}} [C\{\psi_\theta(\boldsymbol{w}x)\} - C\{\psi_{\theta_n}(\boldsymbol{w}x)\}] \frac{dx}{x}.
\end{aligned}$$

As in the treatment of I_2 in the proof of Lemma 4.6, we have that

$$I_{21}(\boldsymbol{w}) \le \sqrt{n} \int_{w_{(1)}/a_n}^{\infty} \frac{1 - \psi_\theta(1/x)}{x} dx \le \sqrt{n} \int_{1/(Ka_n)}^{\infty} \frac{1 - \psi_\theta(1/x)}{x} dx\ .$$

The upper bound is independent of $\boldsymbol{w}$ and converges in probability to 0 by Lemma 5.2 (iii).

To show that $\sup_{\boldsymbol{w}\in B_{1/K}} |I_{22}(\boldsymbol{w})|$ converges to zero in probability, note that $I_{22}(\boldsymbol{w})$ is the same integral as $-\hat{\mathbb{B}}^{\mathrm{CFG}}_{n2}(\boldsymbol{w})$, except for the upper limit of integration. Pick an arbitrary $0 < \delta < x_\Psi/K$; this way, for any $\boldsymbol{w} \in B_{1/K}$, $\delta/w_{(1)} \le x_\Psi/w_j$. Then, for any $\varepsilon > 0$,

$$\Pr\{\sup_{\boldsymbol{w}\in B_{1/K}} |I_{22}(\boldsymbol{w})| > \varepsilon\} = \Pr\{a_n > \delta\} + \Pr\Big\{\sup_{\boldsymbol{w}\in B_{1/K}} \Big|\sqrt{n}\int_0^{\delta/w_{(1)}} [C\{\psi_\theta(\boldsymbol{w}x)\} - C\{\psi_{\theta_n}(\boldsymbol{w}x)\}]\frac{dx}{x}\Big| > \varepsilon\Big\}.$$

The first term on the right-hand side converges to zero because $a_n \to 0$ in probability by Lemma 5.2 (i). As for the second term, the same arguments as in the proof of Lemma 5.5 can then be used to show that

$$\sup_{\boldsymbol{w}\in B_{1/K}} \Big|\int_0^{\delta/w_{(1)}} \sqrt{n}[C\{\psi_\theta(\boldsymbol{w}x)\} - C\{\psi_{\theta_n}(\boldsymbol{w}x)\}]\frac{dx}{x} - a_\delta(\boldsymbol{w})^\top\Theta_n\Big|,$$

converges in probability to 0, where $a_\delta(\boldsymbol{w}) = (a_{\delta,1}(\boldsymbol{w}), \ldots, a_{\delta,p}(\boldsymbol{w}))^\top$ with

$$a_{\delta,k}(\boldsymbol{w}) = -\sum_{j=1}^d \int_0^{\delta/w_{(1)}} \dot{C}_j\{\psi_\theta(\boldsymbol{w}x)\}\dot{\psi}_{\theta,k}(w_j x)\frac{dx}{x}.$$

Observe that as in the proof of Lemma 5.4, for any $k \in \{1, \ldots, p\}$,

$$\begin{aligned}|a_{\delta,k}(\boldsymbol{w})| &\le \sum_{j=1}^d \int_0^{\delta/w_{(1)}} |\dot{\psi}_{\theta,k}(w_j x)|\frac{dx}{x} \\ &\le \sum_{j=1}^d \int_0^{K\delta/w_j} |\dot{\psi}_{\theta,k}(w_j x)|\frac{dx}{x} = d\int_0^{K\delta} |\dot{\psi}_{\theta,k}(t)|\frac{dt}{t} \equiv b_{\delta,k},\end{aligned}$$

so that, using (5.13), $\|b_\delta\|_2 \to 0$ as $\delta \to 0$, where $b_\delta = (b_{\delta,1}, \ldots, b_{\delta,d})^\top$. Hence,

$$\begin{aligned}\limsup_{n\to\infty} \Pr(\sup_{w\in B_{1/K}} |a_\delta(\boldsymbol{w})^\top\Theta_n| > \varepsilon) &\le \limsup_{n\to\infty} \Pr(\|\Theta_n\|_2\|b_\delta\|_2 > \varepsilon) \\ &\le \Pr(\|\Theta\|_2\|b_\delta\|_2 \ge \varepsilon),\end{aligned}$$

where the last inequality is due to the Portmanteau lemma. As $\delta \to 0$, the last expression tends to 0. Put together, we have that $\sup_{\boldsymbol{w}\in B_{1/K}} |I_{22}(\boldsymbol{w})|$ converges in probability to 0, as was to be shown.

Treatment of I_3. Fixing an arbitrary $\boldsymbol{w} \in B_{1/K}$, note that if $x \ge b_n/w_{(d)}$, then $\psi_{\theta_n}(xw_{(d)}) \le c/n < 1/(n+1)$ so that $\hat{C}_n\{\psi_{\theta_n}(\boldsymbol{w}x)\} = 0$. Consequently, $I_3(\boldsymbol{w}) = I_{31}(\boldsymbol{w}) + I_{32}(\boldsymbol{w})$, where

$$\begin{aligned}I_{31}(\boldsymbol{w}) &= \sqrt{n}\int_{b_n/w_{(d)}}^{x_\Psi/w_{(d)}} C\{\psi_\theta(\boldsymbol{w}x)\}\frac{dx}{x} \\ I_{32}(\boldsymbol{w}) &= \sqrt{n}\int_{b_n/w_{(d)}}^{x_\Psi/w_{(d)}} [C\{\psi_{\theta_n}(\boldsymbol{w}x)\} - C\{\psi_\theta(\boldsymbol{w}x)\}]\frac{dx}{x}\ .\end{aligned}$$

As in the treatment of I_3 in the proof of Lemma 4.6,

$$I_{31}(\boldsymbol{w}) \leq \sqrt{n} \int_{b_n/w_{(d)}}^{x_\psi/w_{(d)}} \frac{\psi_\theta(w_{(d)}x)}{x} dx = \sqrt{n} \int_{b_n}^{x_\Psi} \frac{\psi_\theta(x)}{x} dx \,.$$

The upper bound is independent of $\boldsymbol{w}$ and converges in probability to 0 by Lemma 5.2 (ii).

To show that $\sup_{\boldsymbol{w}\in B_{1/K}} |I_{32}(\boldsymbol{w})|$ converges to zero in probability, pick an arbitrary $0 < \kappa < x_\Psi$. Then, for any $\varepsilon > 0$, and κ arbitrarily close to x_Ψ,

$$\Pr\{\sup_{\boldsymbol{w}\in B_{1/K}} |I_{32}(\boldsymbol{w})| > \varepsilon\} = \Pr\{b_n < \kappa\} +$$
$$\Pr\Big\{\sup_{\boldsymbol{w}\in B_{1/K}} \Big|\sqrt{n}\int_{\kappa/w_{(d)}}^{x_\Psi/w_{(d)}} [C\{\psi_{\theta_n}(\boldsymbol{w}x)\} - C\{\psi_\theta(\boldsymbol{w}x)\}]\frac{dx}{x}\Big| > \varepsilon\Big\}.$$

The first term on the right-hand side converges to zero because $b_n \to x_\Psi$ in probability by Lemma 5.2 (i). As for the second term, the same arguments as in the proof of Lemma 5.5 can then be used to show that

$$\sup_{\boldsymbol{w}\in B_{1/K}} \Big|\int_{\kappa/w_{(d)}}^{x_\Psi/w_{(d)}} \sqrt{n}[C\{\psi_\theta(\boldsymbol{w}x)\} - C\{\psi_{\theta_n}(\boldsymbol{w}x)\}]\frac{dx}{x} - a^*_\kappa(\boldsymbol{w})^\top \Theta_n\Big|,$$

converges in probability to 0, where $a^*_\kappa(\boldsymbol{w}) = (a^*_{\kappa,1}(\boldsymbol{w}), \ldots, a^*_{\kappa,p}(\boldsymbol{w}))^\top$ with

$$a^*_{\kappa,k}(\boldsymbol{w}) = \sum_{j=1}^d \int_{\kappa/w_{(d)}}^{x_\Psi/w_{(d)}} \dot{C}_j\{\psi_\theta(\boldsymbol{w}x)\}\dot{\psi}_{\theta,k}(w_j x)\frac{dx}{x}.$$

Because $0 \leq \dot{C}_j \leq 1$, for any $k \in \{1, \ldots, p\}$,

$$|a^*_{\kappa,k}(\boldsymbol{w})| \leq \sum_{j=1}^d \int_{\kappa/w_{(d)}}^{x_\Psi/w_{(d)}} |\dot{\psi}_{\theta,k}(w_j x)|\frac{dx}{x} \,.$$

In the case when $x_\Psi < \infty$, let $M = \sup_{x\in[0,x_\Psi)} \|\dot{\psi}_\theta(x)\|_2$; from (5.14) we have that $M < \infty$. Then $|a^*_{\kappa,k}(\boldsymbol{w})|_2 \leq b^*_{\kappa,k}$, where

$$b^*_{\kappa,k} = dM(\ln x_\Psi - \ln \kappa).$$

Clearly, $\|b^*_\kappa\| \to 0$ as $\kappa \to x_\Psi$, where $b^*_\kappa = (b^*_{\kappa,1}, \ldots, b^*_{\kappa,d})^\top$. In the case when $x_\Psi = \infty$, $|a^*_{\kappa,k}(\boldsymbol{w})| \leq b^*_{\kappa,k}$, where this time,

$$b^*_{\kappa,k} = \sum_{j=1}^d \int_{\kappa/(Kw_j)}^{x_\Psi/w_j} |\dot{\psi}_{\theta,k}(w_j x)|\frac{dx}{x} = d\int_{\kappa/K}^{x_\Psi} |\dot{\psi}_{\theta,k}(t)|\frac{dt}{t},$$

so that, using (5.13), we again have that $\|b^*_\kappa\|_2 \to 0$ as $\kappa \to \infty$, where $b^*_\kappa = (b^*_{\kappa,1}, \ldots, b^*_{\kappa,d})^\top$. Thus when $x_\Psi < \infty$ as well as when $x_\Psi = \infty$,

$$\begin{aligned}\limsup_{n\to\infty} \Pr(\sup_{w\in B_{1/K}} |a^*_\kappa(\boldsymbol{w})^\top \Theta_n| > \varepsilon) &\leq \limsup_{n\to\infty} \Pr(\|\Theta_n\|_2 \|b^*_\kappa\|_2 > \varepsilon) \\ &\leq \Pr(\|\Theta\|_2 \|b^*_\kappa\|_2 \geq \varepsilon),\end{aligned}$$

where the last inequality is due to the Portmanteau lemma. As $\kappa \to x_\Psi$, the upper bound tends to 0, so that $\sup_{\boldsymbol{w}\in B_{1/K}} |I_{32}(\boldsymbol{w})| = o_P(1)$.

Treatment of I_4. Here the second weight function defined for $\boldsymbol{u} \in [0,1]^d$ by $\tilde{g}_\omega(\boldsymbol{u}) = g_\omega(\boldsymbol{u}) + 1\{g_\omega(\boldsymbol{u}) = 0\}$ is used. Letting $\boldsymbol{w} \in B_{1/K}$ and Z_n defined as in (4.16),

$$\begin{aligned} I_4(\boldsymbol{w}) &= \int_0^{a_n/w_{(1)}} \left| \frac{\bar{C}_n\{\psi_{\theta_n}(\boldsymbol{w}x)\}}{\tilde{g}_\omega\{\psi_{\theta_n}(\boldsymbol{w}x)\}} \right| \frac{\tilde{g}_\omega\{\psi_{\theta_n}(\boldsymbol{w}x)\}}{x} dx \\ &\le Z_n \int_0^{a_n/w_{(1)}} \frac{\tilde{g}_\omega\{\psi_{\theta_n}(\boldsymbol{w}x)\}}{x} dx \le Z_n \int_0^{Ka_n} \frac{\tilde{g}_\omega\{\psi_{\theta_n}(\boldsymbol{w}x)\}}{x} dx. \end{aligned}$$

Now suppose for a moment that $a_n \le \delta$ for some δ small enough so that $K\delta < x_\Psi$. Under this assumption,

$$Z_n \int_0^{Ka_n} \frac{\tilde{g}_\omega\{\psi_{\theta_n}(\boldsymbol{w}x)\}}{x} dx = Z_n \int_0^{Ka_n} \frac{g_\omega\{\psi_{\theta_n}(\boldsymbol{w}x)\}}{x} dx.$$

because $g_\omega(\boldsymbol{u}) = 0$ occurs either when at least one component of $\boldsymbol{u}$ equals 0 or at least $d-1$ components equal 1. Write the right-hand side as

$$Z_n \Big[\int_0^{Ka_n} \frac{g_\omega\{\psi_\theta(\boldsymbol{w}x)\}}{x} dx + \int_0^{Ka_n} \frac{g_\omega\{\psi_{\theta_n}(\boldsymbol{w}x)\} - g_\omega\{\psi_\theta(\boldsymbol{w}x)\}}{x} dx \Big]$$

and note from the proof of Lemma 4.6 (Treatment of I_4) that this expression is bounded above by

$$\begin{aligned} &Z_n \int_{1/(K\delta)}^\infty \frac{\{1-\psi_\theta\,(1/x)\}^\omega}{x} dx \\ &\qquad + Z_n \sup_{\boldsymbol{w}\in B_{1/K}} \int_0^{x_\Psi/w_{(d)}} |g_\omega\{\psi_{\theta_n}(\boldsymbol{w}x)\} - g_\omega\{\psi_\theta(\boldsymbol{w}x)\}| \frac{dx}{x}. \end{aligned}$$

Now fix an arbitrary $\varepsilon > 0$ and pick a $\delta > 0$ so that $K\delta < x_\Psi$. Then

$$\Pr(\sup_{\boldsymbol{w}\in B_{1/K}} I_4(\boldsymbol{w}) > \varepsilon) \le \Pr(\sup_{\boldsymbol{w}\in B_{1/K}} I_4(\boldsymbol{w}) > \varepsilon, a_n \le \delta) + \Pr(a_n \ge \delta)$$

Given that $a_n \to 0$ in probability from Lemma 5.2 (i), it suffices to show that the first term on the right-hand side tends to 0 as $n \to \infty$. Write

$$\begin{aligned} &\Pr(\sup_{\boldsymbol{w}\in B_{1/K}} I_4(\boldsymbol{w}) > \varepsilon, a_n \le \delta) \\ &\le \Pr\Big[Z_n \int_{1/(K\delta)}^\infty \frac{\{1-\psi_\theta\,(1/x)\}^\omega}{x} dx > \frac{\varepsilon}{2}\Big] \\ &\quad + \Pr\Big[Z_n \sup_{\boldsymbol{w}\in B_{1/K}} \int_0^{x_\Psi/w_{(d)}} |g_\omega\{\psi_{\theta_n}(\boldsymbol{w}x)\} - g_\omega\{\psi_\theta(\boldsymbol{w}x)\}| \frac{dx}{x} > \frac{\varepsilon}{2}\Big]. \end{aligned}$$

Given that $Z_n \rightsquigarrow Z = \sup_{\boldsymbol{u}\in[0,1]^d} |\mathbb{C}(\boldsymbol{u})/\tilde{g}_\omega(\boldsymbol{u})|$ as $n \to \infty$ by Theorem 2.12, the Portmanteau lemma implies that the lim sup of the first term is bounded above by

$$\Pr\Big[Z \int_{1/(K\delta)}^\infty \frac{\{1-\psi_\theta\,(1/x)\}^\omega}{x} dx \ge \frac{\varepsilon}{2}\Big].$$

This probability can be made arbitrarily small given that

$$\int_{1/(K\delta)}^{\infty} \frac{\{1-\psi_\theta\,(1/x)\}^\omega}{x} dx$$

is bounded above by I_{11} in (4.15), which is finite, and tends to 0 as $\delta \to 0$. Lemma 5.3 and the fact that $Z_n \rightsquigarrow Z$ imply that

$$\lim_{n\to\infty} \Pr\Big[Z_n \sup_{\boldsymbol{w}\in B_{1/K}} \int_0^{x_\Psi/w_{(d)}} |g_\omega\{\psi_{\theta_n}(\boldsymbol{w}x)\} - g_\omega\{\psi_\theta(\boldsymbol{w}x)\}| \frac{dx}{x} > \frac{\varepsilon}{2}\Big] = 0$$

which concludes that $\sup_{\boldsymbol{w}\in B_{1/K}} I_4(\boldsymbol{w}) \to 0$ in probability as $n \to \infty$.

Treatment of I_5. We can proceed similarly as in the preceding paragraph. Fix any $\boldsymbol{w} \in B_{1/K}$ and suppose that $b_n > \delta$ for some $\delta \in (0, x_\Psi)$ arbitrarily close to x_Ψ. Using the arguments from the proof of Lemma 4.6 (treatment of I_5), one has that

$$\begin{aligned} I_5(\boldsymbol{w}) &= \int_{b_n/w_{(d)}}^{x_\Psi/w_{(d)}} \left| \frac{\bar{C}_n\{\psi_{\theta_n}(\boldsymbol{w}x)\}}{\tilde{g}_\omega\{\psi_{\theta_n}(\boldsymbol{w}x)\}} \right| \frac{\tilde{g}_\omega\{\psi_{\theta_n}(\boldsymbol{w}x)\}}{x} dx \\ &\le Z_n \int_{b_n/w_{(d)}}^{x_\Psi/w_{(d)}} \frac{\tilde{g}_\omega\{\psi_{\theta_n}(\boldsymbol{w}x)\}}{x} dx = Z_n \int_{b_n/w_{(d)}}^{x_\Psi/w_{(d)}} \frac{g_\omega\{\psi_{\theta_n}(\boldsymbol{w}x)\}}{x} dx, \end{aligned}$$

and that the upper bound is bounded above by

$$\begin{aligned} &Z_n \int_\delta^{x_\Psi} \frac{\{\psi_\theta(x)\}^\omega}{x} dx \\ &\qquad + Z_n \sup_{\boldsymbol{w}\in B_{1/K}} \int_0^{x_\Psi/w_{(d)}} |g_\omega\{\psi_{\theta_n}(\boldsymbol{w}x)\} - g_\omega\{\psi_\theta(\boldsymbol{w}x)\}| \frac{dx}{x}. \end{aligned}$$

Proceeding as in the proof of $\sup_{\boldsymbol{w}\in B_{1/K}} I_4(\boldsymbol{w}) = o_P(1)$, we thus have that $\sup_{\boldsymbol{w}\in B_{1/K}} I_5(\boldsymbol{w}) = o_P(1)$, since $b_n \to x_\Psi$ in probability as $n \to \infty$ by Lemma 5.2 (i) and $\int_\delta^{x_\Psi} \{\psi_\theta(x)\}^\omega/dx \to 0$ as $\delta \to x_\Psi$ by Lemma 4.4 (i). □

From Lemma 5.6, $\sup_{\boldsymbol{w}\in B_{1/K}} |\hat{\mathbb{B}}_{n1}^{\mathrm{CFG}}(\boldsymbol{w}) - \bar{\mathbb{B}}_{n1}^{\mathrm{CFG}}(\boldsymbol{w})|$ converges to 0 in probability as $n \to \infty$. Finally, introduce $\check{\mathbb{B}}_{n1}^{\mathrm{CFG}}$ given, for all $\boldsymbol{w} \in \Delta_d$, by

$$\check{\mathbb{B}}_{n1}^{\mathrm{CFG}}(\boldsymbol{w}) = -\int_0^{x_\Psi/w_{(d)}} \bar{\mathbb{C}}_n\{\psi_\theta(\boldsymbol{w}x)\} \frac{dx}{x} \tag{5.37}$$

and note the following result.

Lemma 5.7. *As $n \to \infty$, $\sup_{\boldsymbol{w}\in B_{1/K}} |\bar{\mathbb{B}}_{n1}^{CFG}(\boldsymbol{w}) - \check{\mathbb{B}}_{n1}^{CFG}(\boldsymbol{w})|$ converges in probability to 0.*

Proof. Introduce the process $\tilde{\mathbb{B}}_{n1}^{\mathrm{CFG}}$ given, for all $\boldsymbol{w} \in \Delta_d$, by

$$\tilde{\mathbb{B}}_{n1}^{\mathrm{CFG}}(\boldsymbol{w}) = -\int_0^{x_\Psi/w_{(d)}} \frac{\bar{\mathbb{C}}_n\{\psi_{\theta_n}(\boldsymbol{w}x)\}}{\tilde{g}_\omega\{\psi_{\theta_n}(\boldsymbol{w}x)\}} \tilde{g}_\omega\{\psi_\theta(\boldsymbol{w}x)\} \frac{dx}{x}$$

and observe that, with Z_n as in (4.16),

$$\sup_{\boldsymbol{w}\in B_{1/K}} |\bar{\mathbb{B}}_{n1}^{\mathrm{CFG}}(\boldsymbol{w}) - \tilde{\mathbb{B}}_{n1}^{\mathrm{CFG}}(\boldsymbol{w})| \leq$$

$$Z_n \sup_{\boldsymbol{w}\in B_{1/K}} \int_0^{x_\Psi/w_{(d)}} |\tilde{g}_\omega\{\psi_{\theta_n}(\boldsymbol{w}x) - \tilde{g}_\omega\{\psi_\theta(\boldsymbol{w}x)| \frac{dx}{x} .$$

From Theorem 2.12, Z_n converges in law to $\sup_{\boldsymbol{u}\in[0,1]^d} |\mathbb{C}(\boldsymbol{u})/\tilde{g}_\omega(\boldsymbol{u})|$ as $n \to \infty$. Furthermore, because $\phi_\theta(0) = \phi_{\theta_n}(0) = x_\Psi$ from Condition 5.3,

$$\sup_{\boldsymbol{w}\in B_{1/K}} \int_0^{x_\Psi/w_{(d)}} |\tilde{g}_\omega\{\psi_{\theta_n}(\boldsymbol{w}x) - \tilde{g}_\omega\{\psi_\theta(\boldsymbol{w}x)| \frac{dx}{x} =$$

$$\sup_{\boldsymbol{w}\in B_{1/K}} \int_0^{x_\Psi/w_{(d)}} |g_\omega\{\psi_{\theta_n}(\boldsymbol{w}x) - g_\omega\{\psi_\theta(\boldsymbol{w}x)| \frac{dx}{x}.$$

The expression on the right-hand side tends to zero in probability by Lemma 5.3. Consequently, $\sup_{\boldsymbol{w}\in B_{1/K}} |\bar{\mathbb{B}}_{n1}^{\mathrm{CFG}}(\boldsymbol{w}) - \tilde{\mathbb{B}}_{n1}^{\mathrm{CFG}}(\boldsymbol{w})|$ converges to 0 in probability as $n \to \infty$. Next, recall that the sequence (T_n) with T_n as in (5.32) is tight. Hence, for any $\delta > 0$ there exists $M_\delta > 0$ and $N_\delta > 0$ such that for all $n \geq N_\delta$, $\Pr(T_n > M_\delta) < \delta$. Let $\varepsilon > 0$ be arbitrary. Then

$$\Pr(\sup_{\boldsymbol{w}\in B_{1/K}} |\breve{\mathbb{B}}_{n1}^{\mathrm{CFG}}(\boldsymbol{w}) - \bar{\mathbb{B}}_{n1}^{\mathrm{CFG}}(\boldsymbol{w})| > \varepsilon) \leq \delta +$$

$$\Pr\Big\{ \sup_{\substack{\boldsymbol{u},\boldsymbol{u}'\in[0,1]^d \\ \|\boldsymbol{u}-\boldsymbol{u}'\|_2 \leq M_\delta/\sqrt{n}}} \Big| \frac{\bar{\mathbb{C}}_n(\boldsymbol{u})}{\tilde{g}_\omega(\boldsymbol{u})} - \frac{\bar{\mathbb{C}}_n(\boldsymbol{u}')}{\tilde{g}_\omega(\boldsymbol{u}')} \Big| > \frac{\varepsilon}{I_{11} + I_{12}} \Big\},$$

using (4.14); I_{11} and I_{12} are as in (4.15). Because $\delta > 0$ was arbitrary, the conclusion follows from Equation (4.2) of Berghaus et al. (2017). □

Putting all the pieces together, we have that

$$\sup_{\boldsymbol{w}\in B_{1/K}} |\hat{\mathbb{B}}_n^{\mathrm{CFG}}(\boldsymbol{w}) - \breve{\mathbb{B}}_{n1}^{\mathrm{CFG}}(\boldsymbol{w}) - \breve{\mathbb{B}}_{n2}^{\mathrm{CFG}}(\boldsymbol{w})| = o_P(1).$$

Equation (5.12) and the Continuous Mapping Theorem then imply that

$$\hat{\mathbb{B}}_n^{\mathrm{CFG}} \rightsquigarrow -\int_0^{x_\Psi/w_{(d)}} \mathbb{C}\{\psi_\theta(\boldsymbol{w}x)\} \frac{dx}{x} - \sum_{j=1}^d \int_0^{x_\Psi/w_j} \dot{C}_j\{\psi_\theta(\boldsymbol{w}x)\} \dot{\psi}_\theta^\top(w_j x)\Theta \frac{dx}{x}$$

in $\ell^\infty([0,1]^d)$, as was to be shown. The continuity of the mapping follows from (5.30) and the calculations in the last paragraph of Section 4.2.3. Because for any $j \in \{1,\ldots,d\}$, $\dot{C}_j(\boldsymbol{u}) = 0$ if $u_k = 0$ for some $k \neq j$, the limit can be written more succinctly as

$$-\int_0^{x_\Psi/w_{(d)}} \Big[\mathbb{C}\{\psi_\theta(\boldsymbol{w}x)\} + \sum_{j=1}^d \dot{C}_j\{\psi_\theta(\boldsymbol{w}x)\} \dot{\psi}_\theta^\top(w_j x)\Theta\Big] \frac{dx}{x}$$

and by change of variable as

$$\int_0^1 \Big(\mathbb{C}[\psi_\theta\{-\boldsymbol{w}\log(u)\}] + \sum_{j=1}^d \dot{C}_j[\psi_\theta\{-\boldsymbol{w}\log(u)\}]\dot{\psi}_\theta^\top\{-w_j\log(u)\}\Theta\Big)\frac{du}{u\log u}$$

with the convention that, if $x_\Psi < \infty$, $\dot{\psi}_\theta^\top(x) \equiv \mathbf{0}$ whenever $x \geq x_\Psi$.

5.5.4 Proof of Theorem 5.2

The proof proceeds along the same path as the proof of Theorem 5.1. Let $\mathcal{K}$ be a compact subset of $\mathring{\Delta}_d$. For an arbitrary $\boldsymbol{w} \in \Delta_d$, set $w_{(1)} = \min_{i=1,\ldots,d} w_i$ and $w_{(d)} = \max_{i=1,\ldots,d} w_i$. Define, for any $k \in \mathbb{N}$, the set $B_{1/k} = \{\boldsymbol{w} \in \Delta_d : w_{(1)} \geq 1/k\}$. Since $\mathcal{K}$ is compact, there exists an integer $K > 1$ such that $\mathcal{K} \subset B_{1/K} \subset \mathring{\Delta}_d$.

Again, to simplify notation, we denote the true parameter value by θ instead of θ_0 henceforth, and write $\Theta_n = \sqrt{n}(\theta_n - \theta)$.

As in Section 4.1, introduce the process $\hat{\mathbb{B}}_n^{\mathrm{P}}$ given, for all $\boldsymbol{w} \in \Delta_d$, by

$$\hat{\mathbb{B}}_n^{\mathrm{P}}(\boldsymbol{w}) = \sqrt{n}\left\{1/\hat{A}_n^{\mathrm{P}}(\boldsymbol{w}) - 1/A(\boldsymbol{w})\right\}.$$

Proceeding as in the proof of Lemma 4.2, $\hat{\mathbb{B}}_n^{\mathrm{P}}$ may be rewritten as

$$\begin{aligned}\hat{\mathbb{B}}_n^{\mathrm{P}}(\boldsymbol{w}) &= \{\mathrm{E}(Z)\}^{-1}\int_0^\infty \sqrt{n}\left[\hat{C}_n\{\psi_{\theta_n}(\boldsymbol{w}x)\} - C\{\psi_\theta(\boldsymbol{w}x)\}\right]dx \\ &= \{\mathrm{E}(Z)\}^{-1}\int_0^{x_\Psi/w_{(d)}} \sqrt{n}\left[\hat{C}_n\{\psi_{\theta_n}(\boldsymbol{w}x)\} - C\{\psi_\theta(\boldsymbol{w}x)\}\right]dx\end{aligned}$$

where the second equality follows because $\hat{C}_n\{\psi_{\theta_n}(\boldsymbol{w}x)\} = C\{\psi_\theta(\boldsymbol{w}x)\} = 0$ whenever $x > x_\Psi/w_{(d)}$ if Condition 5.3 holds. Next, write $\hat{\mathbb{B}}_n^{\mathrm{P}} = \hat{\mathbb{B}}_{n1}^{\mathrm{P}} + \hat{\mathbb{B}}_{n2}^{\mathrm{P}}$, where for all $\boldsymbol{w} \in \Delta_d$,

$$\hat{\mathbb{B}}_{n1}^{\mathrm{P}}(\boldsymbol{w}) = \{\mathrm{E}(Z)\}^{-1}\int_0^{x_\Psi/w_{(d)}} \hat{\mathbb{C}}_n\{\psi_{\theta_n}(\boldsymbol{w}x)\}dx$$

and

$$\hat{\mathbb{B}}_{n2}^{\mathrm{P}}(\boldsymbol{w}) = \{\mathrm{E}(Z)\}^{-1}\int_0^{x_\Psi/w_{(d)}} \sqrt{n}\left[C\{\psi_{\theta_n}(\boldsymbol{w}x)\} - C\{\psi_\theta(\boldsymbol{w}x)\}\right]dx\,.$$

As in the proof of Theorem 5.1, it is important to establish weak convergence of the drift $\hat{\mathbb{B}}_{n2}^{\mathrm{P}}$ first. To this end, let $\check{\mathbb{B}}_{n2}^{\mathrm{P}}$ be the process given for all $\boldsymbol{w} \in \mathring{\Delta}_d$ by $\check{\mathbb{B}}_{n2}^{\mathrm{P}}(\boldsymbol{w}) = a^\top(\boldsymbol{w})\Theta_n$ where $a(\boldsymbol{w}) = (a_1(\boldsymbol{w}), \ldots, a_p(\boldsymbol{w}))^\top$ with

$$a_k(\boldsymbol{w}) = \{\mathrm{E}(Z)\}^{-1}\sum_{j=1}^d \int_0^{x_\Psi/w_j} \dot{C}_j\{\psi_\theta(\boldsymbol{w}x)\}\dot{\psi}_{\theta,k}(w_jx)dx. \tag{5.38}$$

The following lemma shows that $|a_k(\boldsymbol{w})| < \infty$ for any $k \in \{1, \ldots, p\}$ and $\boldsymbol{w} \in \mathring{\Delta}_d$, and determines the asymptotic behavior of $\check{\mathbb{B}}_{n2}^{\mathrm{P}}$.

Lemma 5.8. *As $n \to \infty$, $\check{\mathbb{B}}_{n2}^{\mathrm{P}} \rightsquigarrow \mathbb{B}_2^{\mathrm{P}}$ in $\mathcal{C}(B_{1/K})$, where for all $\boldsymbol{w} \in \mathring{\Delta}_d$, $\mathbb{B}_2^{\mathrm{P}}(\boldsymbol{w}) = a^\top(\boldsymbol{w})\Theta$.*

Proof. Fix an arbitrary $K \geq 2$. and note that for any $k \in \{1, \ldots, p\}$ and $\boldsymbol{w} \in \mathring{\Delta}_d$, $|a_k(\boldsymbol{w})| < \infty$. Indeed, since $0 \leq \dot{C}_j \leq 1$ for all $j \in \{1, \ldots, d\}$, we have that

$$\begin{aligned} \mathrm{E}(Z)|a_k(\boldsymbol{w})| &\leq \sum_{j=1}^{d} \int_0^{x_\Psi/w_j} \dot{C}_j\{\psi_\theta(\boldsymbol{w}x)\}|\dot{\psi}_{\theta,k}(w_j x)|dx \\ &\leq \sum_{j=1}^{d} \int_0^{x_\Psi/w_j} |\dot{\psi}_{\theta,k}(w_j x)|dx = \sum_{j=1}^{d} \frac{1}{w_j} \int_0^{x_\Psi} |\dot{\psi}_{\theta,k}(t)|dt \end{aligned}$$

The last expression is finite by Condition 5.5. Next we show that a is uniformly continuous on $B_{1/K}$ viz.

$$\lim_{\delta \downarrow 0} \sup_{\boldsymbol{w}, \boldsymbol{w}' \in B_{1/K}, \|\boldsymbol{w}-\boldsymbol{w}'\| < \delta} \|a(\boldsymbol{w}) - a(\boldsymbol{w}')\|_2 = 0. \tag{5.39}$$

To show that (5.39) holds, define, for all $j \in \{1, \ldots, d\}$ and $k \in \{1, \ldots, p\}$,

$$b_{j,k}(\boldsymbol{w}) = \int_0^{x_\Psi/w_j} \dot{C}_j\{\psi_\theta(\boldsymbol{w}x)\}\dot{\psi}_{\theta,k}(w_j x)dx.$$

Then (5.39) follows if for all $j \in \{1, \ldots, d\}$ and $k \in \{1, \ldots, p\}$,

$$\lim_{\delta \downarrow 0} \sup_{\boldsymbol{w}, \boldsymbol{w}' \in B_{1/K}, \|\boldsymbol{w}-\boldsymbol{w}'\|_2 < \delta} |b_{j,k}(\boldsymbol{w}) - b_{j,k}(\boldsymbol{w}')| = 0.$$

Pick an arbitrary $j \in \{1, \ldots, d\}$, $k \in \{1, \ldots, p\}$. Then for any $\boldsymbol{w}, \boldsymbol{w}' \in B_{1/K}$,

$$\begin{aligned} &|b_{j,k}(\boldsymbol{w}) - b_{j,k}(\boldsymbol{w}')| \\ &\quad \leq \int_0^{x_\Psi} \left| \frac{\dot{C}_j\{\psi_\theta(\boldsymbol{w}t/w_j)\}}{w_j} - \frac{\dot{C}_j\{\psi_\theta(\boldsymbol{w}'t/w'_j)\}}{w'_j} \right| |\dot{\psi}_{\theta,k}(t)|dt \\ &\quad \leq \int_0^{x_\Psi} \frac{|w'_j \dot{C}_j\{\psi_\theta(\boldsymbol{w}t/w_j)\} - w_j \dot{C}_j\{\psi_\theta(\boldsymbol{w}'t/w'_j)\}|}{w_j w'_j} |\dot{\psi}_{\theta,k}(t)|dt \\ &\quad \leq K^2 w_j \int_0^{x_\Psi} |\dot{C}_j\{\psi_\theta(\boldsymbol{w}t/w_j)\} - \dot{C}_j\{\psi_\theta(\boldsymbol{w}'t/w'_j)\}||\dot{\psi}_{\theta,k}(t)|dt \\ &\qquad + K^2|w_j - w'_j| \int_0^{x_\Psi} \dot{C}_j\{\psi_\theta(\boldsymbol{w}t/w_j)\}|\dot{\psi}_{\theta,k}(t)|dt \\ &\quad \leq K^2 \int_0^{x_\Psi} |\dot{C}_j\{\psi_\theta(\boldsymbol{w}t/w_j)\} - \dot{C}_j\{\psi_\theta(\boldsymbol{w}'t/w'_j)\}||\dot{\psi}_{\theta,k}(t)|dt \\ &\qquad + K^2|w_j - w'_j| \int_0^{x_\Psi} |\dot{\psi}_{\theta,k}(t)|dt\ . \end{aligned}$$

Due to the fact that $\int_0^{x_\Psi} |\dot{\psi}_{\theta,k}(t)|dt$ is finite by (5.17),

$$\sup_{\substack{\boldsymbol{w}, \boldsymbol{w}' \in B_{1/K}, \\ \|\boldsymbol{w}-\boldsymbol{w}'\|_2 < \delta}} K^2|w_j - w'_j| \int_0^{x_\Psi} |\dot{\psi}_{\theta,k}(t)|dt \to 0$$

as $\delta \to 0$. The rest of the argument follows as in the proof of Lemma 5.4. Pick an arbitrary $\eta, \mu \in (0, 1)$ and note that because ψ_θ is uniformly continuous, there exists

$\lambda > 0$ such that for all $|x-x'| < \lambda$, $|\psi_\theta(x) - \psi_\theta(x')| < \mu$. Also note that if $\|\boldsymbol{w}-\boldsymbol{w}'\|_2 < \delta$, $|(w_k t/w_j) - (w'_k t/w'_j)| \le 2K^2 t\delta$. Because $2K^2\phi_\theta(\eta)\delta < \lambda$ for all δ sufficiently small and because $0 \le \dot{C}_j \le 1$,

$$\lim_{\delta\downarrow 0} \sup_{\substack{\boldsymbol{w},\boldsymbol{w}'\in B_{1/K},\\ \|\boldsymbol{w}-\boldsymbol{w}'\|_2<\delta}} |b_{j,k}(\boldsymbol{w}) - b_{j,k}(\boldsymbol{w}')| \le$$

$$K^2 \sup_{\substack{\boldsymbol{u},\boldsymbol{u}'\in A_{\eta,j}\\ \|\boldsymbol{u}-\boldsymbol{u}'\|_2<\mu}} |\dot{C}_j(\boldsymbol{u}) - \dot{C}_j(\boldsymbol{u}')| \int_{\phi_\theta(1-\eta)}^{\phi_\theta(\eta)} |\dot{\psi}_{\theta,k}(t)|dt$$

$$+ 2K^2\Big\{\int_0^{\phi_\theta(1-\eta)} |\dot{\psi}_{\theta,k}(t)|dt + \int_{\phi_\theta(\eta)}^{x_\Psi} |\dot{\psi}_{\theta,k}(t)|dt\Big\},$$

where $A_{\eta,j} = \{\boldsymbol{u} \in [0,1]^d : u_j \in [\eta, 1-\eta]\}$. Because $\dot{C}_j$ is uniformly continuous on the set $A_{\eta,j}$ by Proposition 3.2, the first expression on the right-hand side tends to 0 as $\mu \to 0$. Because $\int_0^{x_\Psi} |\dot{\psi}_{\theta,k}(t)|dt$ is finite by Condition 5.5, the second expression tends to 0 as $\eta \to 0$. □

The following result shows that $\hat{\mathbb{B}}^{\mathrm{P}}_{n2}$ behaves asymptotically as $\check{\mathbb{B}}^{\mathrm{P}}_{n2}$.

Lemma 5.9. *As $n \to \infty$, $\sup_{\boldsymbol{w}\in B_{1/K}} |\check{\mathbb{B}}^{\mathrm{P}}_{n2}(\boldsymbol{w}) - \hat{\mathbb{B}}^{\mathrm{P}}_{n2}(\boldsymbol{w})|$ converges to 0 in probability.*

Proof. Let $\tilde{\mathbb{B}}^{\mathrm{P}}_{n2}$ be given, for all $\boldsymbol{w} \in \mathring{\Delta}_d$, by

$$\tilde{\mathbb{B}}^{\mathrm{P}}_{n2}(\boldsymbol{w}) = \{\mathrm{E}(Z)\}^{-1} \sum_{j=1}^d \int_0^{x_\Psi/w_j} \sqrt{n}\{\psi_{\theta_n}(w_j x) - \psi_\theta(w_j x)\}\dot{C}_j\{\psi_\theta(\boldsymbol{w}x)\}dx.$$

We will first show that

$$\sup_{\boldsymbol{w}\in B_{1/K}} |\tilde{\mathbb{B}}^{\mathrm{P}}_{n2}(\boldsymbol{w}) - \hat{\mathbb{B}}^{\mathrm{P}}_{n2}(\boldsymbol{w})| = o_P(1). \tag{5.40}$$

To this end, use the Mean-Value Theorem to write

$$\begin{aligned}\hat{\mathbb{B}}^{\mathrm{P}}_{n2}(\boldsymbol{w}) &= \{\mathrm{E}(Z)\}^{-1} \int_0^\infty \sqrt{n}\,[C\{\psi_{\theta_n}(\boldsymbol{w}x)\} - C\{\psi_\theta(\boldsymbol{w}x)\}]\,dx \\ &= \{\mathrm{E}(Z)\}^{-1} \sum_{j=1}^d \int_0^{x_\Psi/w_j} \sqrt{n}\{\psi_{\theta_n}(w_j x) - \psi_\theta(w_j x)\}\dot{C}_j(\boldsymbol{u}_{\boldsymbol{w}x})dx,\end{aligned}$$

where for every $\boldsymbol{w}x$ and realization ϖ, $\boldsymbol{u}_{\boldsymbol{w}x}(\varpi) = \epsilon(\boldsymbol{w}x,\varpi)\psi_{\theta_n(\varpi)}(\boldsymbol{w}x) + \{1-\epsilon(t,\varpi)\}\psi_\theta(\boldsymbol{w}x)$ for some $\epsilon(\boldsymbol{w}x,\varpi) \in [0,1]$. It thus suffices to show that for all $j \in \{1,\dots,d\}$,

$$V_n = \sup_{\boldsymbol{w}\in B_{1/K}} \left| \int_0^{x_\Psi/w_j} \sqrt{n}\{\psi_{\theta_n}(w_j x) - \psi_\theta(w_j x)\}[\dot{C}_j\{\psi_\theta(\boldsymbol{w}x)\} - \dot{C}_j(\boldsymbol{u}_{\boldsymbol{w}x})]dx \right|$$

converges in probability to 0 as $n \to \infty$. To accomplish this, fix an arbitrary $j \in \{1,\dots,d\}$ and let T_n be defined as in (5.32). From Lemma 5.1 (i), it follows that the sequence (T_n)

is tight. For any $\delta > 0$ there exists $M_\delta > 0$ and $N_\delta > 0$ such that for all $n \geq N_\delta$, $\Pr(T_n > M_\delta) < \delta$. Pick an arbitrary $\varepsilon > 0$, $\eta \in (0,1)$ and let $n \geq N_\delta$ be such that $M_\delta/\sqrt{n} < \eta/2$. Then $\Pr(V_n > \varepsilon) \leq \delta + \Pr(V_n > \varepsilon, T_n \leq M_\delta)$ and $\Pr(V_n > \varepsilon, T_n \leq M_\delta)$ may be bounded above by $\Pr(V_{n1} > \varepsilon/2) + \Pr(V_{n2} > \varepsilon/2)$, where analogously to the proof of Lemma 5.5,

$$\begin{aligned}
V_{n1} &= 2 \sup_{\boldsymbol{w} \in B_{1/K}} \Big\{ \int_0^{\phi_\theta(1-\eta)/w_j} \sqrt{n} |\psi_{\theta_n}(w_j x) - \psi_\theta(w_j x)| dx \\
&\quad + \int_{\phi_\theta(\eta)/w_j}^{x_\Psi/w_j} \sqrt{n} |\psi_{\theta_n}(w_j x) - \psi_\theta(w_j x)| dx \Big\} \\
&= 2K \Big\{ \int_0^{\phi_\theta(1-\eta)} \sqrt{n} |\psi_{\theta_n}(t) - \psi_\theta(t)| dt + \int_{\phi_\theta(\eta)}^{x_\Psi} \sqrt{n} |\psi_{\theta_n}(t) - \psi_\theta(t)| dt \Big\}
\end{aligned}$$

and

$$V_{n2} = \sup_{\substack{\boldsymbol{u}, \boldsymbol{u}' \in A_{\eta/2,j} \\ \|\boldsymbol{u} - \boldsymbol{u}'\|_2 < M_\delta/\sqrt{n}}} |\dot{C}_j(\boldsymbol{u}) - \dot{C}_j(\boldsymbol{u}')| K \int_{\phi_\theta(1-\eta)}^{\phi_\theta(\eta)} \sqrt{n} |\psi_{\theta_n}(t) - \psi_\theta(t)| dt,$$

where $A_{\eta/2,j} = \{\boldsymbol{u} \in [0,1]^d : u_j \in [\eta/2, 1-\eta/2]\}$. Because $\dot{C}_j$ is uniformly continuous on $A_{\eta/2,j}$ and

$$\int_{\phi_\theta(1-\eta)}^{\phi_\theta(\eta)} \sqrt{n} |\psi_{\theta_n}(x) - \psi_\theta(x)| dx \rightsquigarrow \int_{\phi_\theta(1-\eta)}^{\phi_\theta(\eta)} |\dot{\psi}_\theta^\top(x) \Theta| dx$$

as $n \to \infty$ by Lemma 5.1 (iv), $V_{n2} \to 0$ in probability as $n \to \infty$. The same lemma, again part (iv), also implies that as $n \to \infty$,

$$V_{n1} \rightsquigarrow 2K \Big\{ \int_0^{\phi_\theta(1-\eta)} |\dot{\psi}_\theta^\top(x) \Theta| dx + \int_{\phi_\theta(\eta)}^{x_\Psi} |\dot{\psi}_\theta^\top(x) \Theta| dx \Big\}.$$

The limit is non-negative and bounded above by

$$2K \sup_{t \in [0, x_\Psi)} \frac{|\dot{\psi}_\theta^\top(t) \Theta|}{h_\theta(t)} \Big\{ \int_0^{\phi_\theta(1-\eta)} h_\theta(x) dx + \int_{\phi_\theta(\eta)}^{x_\Psi} h_\theta(x) dx \Big\}$$

By the Portmanteau lemma, $\limsup_{n\to\infty} \Pr(V_{n1} > \varepsilon/2)$ is at most

$$\Pr\Big[2K \sup_{t \in [0, x_\Psi)} \frac{|\dot{\psi}_\theta^\top(t) \Theta|}{h_\theta(t)} \Big\{ \int_0^{\phi_\theta(1-\eta)} h_\theta(x) dx + \int_{\phi_\theta(\eta)}^{x_\Psi} h_\theta(x) dx \Big\} \geq \varepsilon/2 \Big].$$

This probability can be made arbitrarily small given that

$$\lim_{\eta \to 0} \Big\{ \int_0^{\phi_\theta(1-\eta)} h_\theta(x) dx + \int_{\phi_\theta(\eta)}^{x_\Psi} h_\theta(x) dx \Big\} = 0.$$

Since δ was arbitrary, $\Pr(V_n > \varepsilon) \to 0$ as $n \to \infty$. This establishes (5.40).

Next, we will prove that

$$\sup_{\boldsymbol{w} \in B_{1/K}} |\tilde{\mathbb{B}}_{n2}^{\mathrm{P}}(\boldsymbol{w}) - \check{\mathbb{B}}_{n2}^{\mathrm{P}}(\boldsymbol{w})| = o_P(1). \tag{5.41}$$

To this end, it suffices to show that for each $j \in \{1, \dots, d\}$,

$$\sup_{\boldsymbol{w} \in B_{1/K}} \left| \int_0^{x_\Psi/w_j} \dot{C}_j\{\psi_\theta(\boldsymbol{w}x)\}[\dot{\psi}_\theta^\top(w_jx)\Theta_n - \sqrt{n}\{\psi_{\theta_n}(w_jx) - \psi_\theta(w_jx)\}]dx \right|$$

converges to 0 in probability as $n \to \infty$. Given that this expression is bounded above by

$$W_n = K \int_0^{x_\Psi} |\dot{\psi}_\theta^\top(t)\Theta_n - \sqrt{n}\{\psi_{\theta_n}(t) - \psi_\theta(t)\}|dt,$$

one can proceed as when showing (5.33) in the proof of Lemma 5.5. □

From Lemmas 5.8 and 5.9,

$$\hat{\mathbb{B}}_{n2}^{\mathrm{P}} \rightsquigarrow \mathbb{B}_2^{\mathrm{P}} \tag{5.42}$$

as $n \to \infty$ in $\mathcal{C}(B_{1/K})$, where for all $\boldsymbol{w} \in \mathring{\Delta}_d$, $\mathbb{B}_2^{\mathrm{P}}(\boldsymbol{w}) = a^\top(\boldsymbol{w})\Theta$. Next, let $\bar{\mathbb{C}}_n$ be as in Theorem 2.12 in Section 4.2.3 and define for all $\boldsymbol{w} \in \Delta_d$,

$$\bar{\mathbb{B}}_{n1}^{\mathrm{P}}(\boldsymbol{w}) = \{\mathrm{E}(Z)\}^{-1} \int_0^{x_\Psi/w_{(d)}} \bar{\mathbb{C}}_n\{\psi_{\theta_n}(\boldsymbol{w}x)\}dx,$$

where $\bar{\mathbb{C}}_n$ is as defined in Theorem 2.12. The following result is the analogue of Lemma 4.7.

Lemma 5.10. *As $n \to \infty$,*

$$\sup_{\boldsymbol{w} \in B_{1/K}} \int_0^{x_\Psi/w_{(d)}} |\hat{\mathbb{C}}_n\{\psi_{\theta_n}(\boldsymbol{w}x)\} - \bar{\mathbb{C}}_n\{\psi_{\theta_n}(\boldsymbol{w}x)\}|dx$$

converges to 0 *in probability.*

Proof. Fix $\omega \in (0, 1/2)$; if Condition 3.2 (a) holds, it is also required that $s\omega > 1$. Define the sequences a_n and b_n and the constant N_K as in Lemma 5.6 and fix $c \in (0, 1/K)$. Then,

$$\int_0^{x_\Psi/w_{(d)}} |\hat{\mathbb{C}}_n\{\psi_{\theta_n}(\boldsymbol{w}x)\} - \bar{\mathbb{C}}_n\{\psi_{\theta_n}(\boldsymbol{w}x)\}|dx \leq \sum_{j=1}^{5} I_j(\boldsymbol{w}),$$

where

$$I_1(\boldsymbol{w}) = \int_{a_n/w_{(1)}}^{b_n/w(d)} \left| \hat{\mathbb{C}}_n\{\psi_{\theta_n}(\boldsymbol{w}x) - \bar{\mathbb{C}}_n\{\psi_{\theta_n}(\boldsymbol{w}x) \right| dx,$$

$$I_2(\boldsymbol{w}) = \int_0^{a_n/w_{(1)}} \left| \hat{\mathbb{C}}_n\{\psi_{\theta_n}(\boldsymbol{w}x) \right| dx, \; I_3(\boldsymbol{w}) = \int_{b_n/w(d)}^{x_\Psi/w_{(d)}} \left| \hat{\mathbb{C}}_n\{\psi_{\theta_n}(\boldsymbol{w}x) \right| dx,$$

$$I_4(\boldsymbol{w}) = \int_0^{a_n/w_{(1)}} \left| \bar{\mathbb{C}}_n\{\psi_{\theta_n}(\boldsymbol{w}x) \right| dx, \; I_5(\boldsymbol{w}) = \int_{b_n/w(d)}^{x_\Psi/w_{(d)}} \left| \bar{\mathbb{C}}_n\{\psi_{\theta_n}(\boldsymbol{w}x) \right| dx\ .$$

Next, each integral is shown to converge to zero in probability as $n \to \infty$.

Treatment of I_1. With S_n is as in (4.11), for any $\boldsymbol{w} \in B_{1/K}$,

$$I_1(\boldsymbol{w}) \leq S_n \left[\int_0^{x_\Psi/w_{(d)}} g_\omega\{\psi_\theta(\boldsymbol{w}x)\}dx \right.$$
$$\left. + \sup_{\boldsymbol{w}\in B_{1/K}} \int_0^{x_\Psi/w_{(d)}} |g_\omega\{\psi_{\theta_n}(\boldsymbol{w}x)\} - g_\omega\{\psi_\theta(\boldsymbol{w}x)\}|dx \right].$$

By the first part of Theorem 2.12, S_n converges to 0 in probability as $n \to \infty$, while Lemma 5.3 (ii) implies that the term in the square brackets converges in probability to

$$\int_0^{x_\Psi/w_{(d)}} g_\omega\{\psi_\theta(\boldsymbol{w}x)\}dx,$$

which was shown to be finite while discussing I_1 in the proof of Lemma 4.7.

Treatment of I_2. Fix $\boldsymbol{w} \in B_{1/K}$. Similarly to the treatment of I_2 in the proof of Lemma 5.6, $I_2(\boldsymbol{w}) = I_{21}(\boldsymbol{w}) + I_{22}(\boldsymbol{w})$ where

$$I_{21}(\boldsymbol{w}) = \sqrt{n} \int_0^{a_n/w_{(1)}} [1 - C\{\psi_\theta(\boldsymbol{w}x)\}]dx,$$
$$I_{22}(\boldsymbol{w}) = \sqrt{n} \int_0^{a_n/w_{(1)}} [C\{\psi_\theta(\boldsymbol{w}x)\} - C\{\psi_{\theta_n}(\boldsymbol{w}x)\}]dx.$$

Since $I_{21}(\boldsymbol{w}) \leq \sqrt{n}\int_{1/(Ka_n)}^\infty \{1-\psi(1/x)\}/x^2dx$, Lemma 5.2 (iv) ensures convergence to zero in probability, uniformly on $B_{1/K}$. The second integral I_{22} is the same as $\mathrm{E}(Z)\hat{\mathbb{B}}_n^{\mathrm{P}}$ but with a different upper limit of integration. Fix an arbitrary $\delta \in (0, x_\Psi/K)$ so that for all $\boldsymbol{w} \in B_{1/K}$, $\delta/w_{(1)} \leq x_\Psi/w_j$ for all $j = 1, \ldots, d$. Then for any $\varepsilon > 0$,

$$\Pr\{\sup_{\boldsymbol{w}\in B_{1/K}} |I_{22}(\boldsymbol{w})| > \varepsilon\} = \Pr\{a_n > \delta\} +$$
$$\Pr\Big\{\sup_{\boldsymbol{w}\in B_{1/K}} \Big|\sqrt{n}\int_0^{\delta/w_{(1)}} [C\{\psi_\theta(\boldsymbol{w}x)\} - C\{\psi_{\theta_n}(\boldsymbol{w}x)\}]\Big|dx > \varepsilon\Big\}.$$

By Lemma 5.2 (i), $\Pr\{a_n > \delta\} \to 0$ as $n \to \infty$. The same approach as in the proof of Lemma 5.9 can then be used to show that

$$\sup_{\boldsymbol{w}\in B_{1/K}} \Big| \int_0^{\delta/w_{(1)}} \sqrt{n}[C\{\psi_\theta(\boldsymbol{w}x)\} - C\{\psi_{\theta_n}(\boldsymbol{w}x)\}]dx - a_\delta(\boldsymbol{w})^\top \Theta_n \Big|,$$

converges in probability to 0, where $a_\delta(\boldsymbol{w}) = (a_{\delta,1}(\boldsymbol{w}), \ldots, a_{\delta,p}(\boldsymbol{w}))^\top$ with

$$a_{\delta,k}(\boldsymbol{w}) = \sum_{j=1}^d \int_0^{\delta/w_{(1)}} \dot{C}_j\{\psi_\theta(\boldsymbol{w}x)\}\dot{\psi}_{\theta,k}(w_jx)dx.$$

Analogously to the proof of Lemma 5.8, for any $k \in \{1, \ldots, p\}$,

$$|a_{\delta,k}(\boldsymbol{w})| \leq \sum_{j=1}^d \int_0^{\delta/w_{(1)}} |\dot{\psi}_{\theta,k}(w_jx)|dx$$
$$\leq \sum_{j=1}^d \int_0^{K\delta/w_j} |\dot{\psi}_{\theta,k}(w_jx)|dx \leq Kd\int_0^{K\delta} |\dot{\psi}_{\theta,k}(t)|dt \equiv b_{\delta,k},$$

and using (5.17), $\|b_\delta\|_2 \to 0$ as $\delta \to 0$, where $b_\delta = (b_{\delta,1}, \ldots, b_{\delta,d})^\top$. Hence,

$$\limsup_{n\to\infty} \Pr(\sup_{w\in B_{1/K}} |a_\delta(\boldsymbol{w})^\top \Theta_n| > \varepsilon) \le \limsup_{n\to\infty} \Pr(\|\Theta_n\|_2 \|b_\delta\|_2 > \varepsilon)$$
$$\le \Pr(\|\Theta\|_2 \|b_\delta\|_2 \ge \varepsilon),$$

where the last inequality is due to the Portmanteau lemma. As $\delta \to 0$, the last expression tends to 0. We can conclude that $\sup_{\boldsymbol{w}\in B_{1/K}} |I_{22}(\boldsymbol{w})|$ converges in probability to 0, as needed.

Treatment of I_3. For any $\boldsymbol{w} \in B_{1/K}$, $I_3(\boldsymbol{w}) = I_{31}(\boldsymbol{w}) + I_{32}(\boldsymbol{w})$, where

$$I_{31}(\boldsymbol{w}) = \sqrt{n} \int_{b_n/w_{(d)}}^{x_\Psi/w_{(d)}} C\{\psi_\theta(\boldsymbol{w}x)\}dx$$
$$I_{32}(\boldsymbol{w}) = \sqrt{n} \int_{b_n/w_{(d)}}^{x_\Psi/w_{(d)}} [C\{\psi_{\theta_n}(\boldsymbol{w}x)\} - C\{\psi_\theta(\boldsymbol{w}x)\}]dx\ .$$

As in the treatment of I_3 in the proof of Lemma 4.7,

$$I_{31}(\boldsymbol{w}) \le \sqrt{n} \int_{b_n/w_{(d)}}^{x_\psi/w_{(d)}} \psi_\theta(w_{(d)}x)dx \le K\sqrt{n} \int_{b_n}^{x_\Psi} \psi(x)dx\ .$$

By Lemma 5.2 (iii), the upper bound converges in probability to 0.

Now pick an arbitrary $\kappa \in (0, x_\Psi)$. Then, for any $\varepsilon > 0$, and κ arbitrarily close to x_Ψ,

$$\Pr\{\sup_{\boldsymbol{w}\in B_{1/K}} |I_{32}(\boldsymbol{w})| > \varepsilon\} = \Pr\{b_n < \kappa\} +$$
$$\Pr\Big\{\sup_{\boldsymbol{w}\in B_{1/K}} \Big|\sqrt{n} \int_{\kappa/w_{(d)}}^{x_\Psi/w_{(d)}} [C\{\psi_{\theta_n}(\boldsymbol{w}x)\} - C\{\psi_\theta(\boldsymbol{w}x)\}]\Big| dx > \varepsilon\Big\}.$$

By Lemma 5.2 (i), the first term on the right-hand side converges to zero. For the second term, the same arguments as in the proof of Lemma 5.9 can then be used to show that

$$\sup_{\boldsymbol{w}\in B_{1/K}} \Big| \int_{\kappa/w_{(d)}}^{x_\Psi/w_{(d)}} \sqrt{n}[C\{\psi_\theta(\boldsymbol{w}x)\} - C\{\psi_{\theta_n}(\boldsymbol{w}x)\}]dx - a^*_\kappa(\boldsymbol{w})^\top \Theta_n \Big|,$$

converges in probability to 0, where $a^*_\kappa(\boldsymbol{w}) = (a^*_{\kappa,1}(\boldsymbol{w}), \ldots, a^*_{\kappa,p}(\boldsymbol{w}))^\top$ with

$$a^*_{\kappa,k}(\boldsymbol{w}) = \sum_{j=1}^d \int_{\kappa/w_{(d)}}^{x_\Psi/w_{(d)}} \dot{C}_j\{\psi_\theta(\boldsymbol{w}x)\}\dot{\psi}_{\theta,k}(w_j x)dx.$$

Since for any $k \in \{1, \ldots, p\}$ and $\boldsymbol{u} \in [0,1]^d$, $\dot{C}_j(\boldsymbol{u}) \in [0,1]$, we have that

$$|a^*_{\kappa,k}(\boldsymbol{w})| \le \sum_{j=1}^d \int_{\kappa/w_{(d)}}^{x_\Psi/w_{(d)}} |\dot{\psi}_{\theta,k}(w_j x)|dx\ .$$

In the case when $x_\Psi < \infty$, let $M = \sup_{x\in[0,x_\Psi)} \|\dot\psi_\theta(x)\|_2$; from (5.14) we have that $M < \infty$. Then $|a^*_{\kappa,k}(\boldsymbol{w})| \le MK(x_\Psi - \kappa) \equiv b^*_{\kappa,k}$. Clearly, $\|b^*_\kappa\|_2 \to 0$ as $\kappa \to x_\Psi$, where $b^*_\kappa = (b^*_{\kappa,1},\ldots,b^*_{\kappa,d})^\top$. If $x_\Psi = \infty$, $|a^*_{\kappa,k}(\boldsymbol{w})| \le b^*_{\kappa,k}$ with

$$b^*_{\kappa,k} = \sum_{j=1}^{d} \int_{\kappa/(Kw_j)}^{x_\Psi/w_j} |\dot\psi_{\theta,k}(w_j x)| dx \le dK \int_{\kappa/K}^{x_\Psi} |\dot\psi_{\theta,k}(t)| dt,$$

so that, using (5.17), we again have that $\|b^*_\kappa\|_2 \to 0$ as $\kappa \to \infty$, where $b^*_\kappa = (b^*_{\kappa,1},\ldots,b^*_{\kappa,d})^\top$. Hence,

$$\begin{aligned}\limsup_{n\to\infty} \Pr(\sup_{w\in B_{1/K}} |a^*_\kappa(\boldsymbol{w})^\top \Theta_n| > \varepsilon) &\le \limsup_{n\to\infty} \Pr(\|\Theta_n\|_2 \|b^*_\kappa\|_2 > \varepsilon) \\ &\le \Pr(\|\Theta\|_2 \|b^*_\kappa\|_2 \ge \varepsilon),\end{aligned}$$

where the last inequality is due to the Portmanteau lemma. As $\kappa \to x_\Psi$, the upper bound tends to 0, so that $\sup_{\boldsymbol{w}\in B_{1/K}} |I_{32}(\boldsymbol{w})| = o_P(1)$.

Treatment of I_4. Recall that for $\boldsymbol{u} \in [0,1]^d$, $\tilde g_\omega(\boldsymbol{u}) = g_\omega(\boldsymbol{u}) + 1\{g_\omega(\boldsymbol{u}) = 0\}$. Letting $\boldsymbol{w} \in B_{1/K}$ and Z_n defined as in (4.16),

$$\begin{aligned} I_4(\boldsymbol{w}) &= \int_0^{a_n/w_{(1)}} \left| \frac{\bar C_n\{\psi_{\theta_n}(\boldsymbol{w}x)\}}{\tilde g_\omega\{\psi_{\theta_n}(\boldsymbol{w}x)\}} \right| \tilde g_\omega\{\psi_{\theta_n}(\boldsymbol{w}x)\} dx \\ &\le Z_n \int_0^{a_n/w_{(1)}} \tilde g_\omega\{\psi_{\theta_n}(\boldsymbol{w}x)\} dx \le Z_n \int_0^{Ka_n} \tilde g_\omega\{\psi_{\theta_n}(\boldsymbol{w}x)\} dx. \end{aligned}$$

Suppose that $a_n \le \delta$ for some δ small enough so that $K\delta < x_\Psi$. Then

$$Z_n \int_0^{Ka_n} \tilde g_\omega\{\psi_{\theta_n}(\boldsymbol{w}x)\} dx = Z_n \int_0^{Ka_n} g_\omega\{\psi_{\theta_n}(\boldsymbol{w}x)\} dx$$

because $g_\omega(\boldsymbol{u}) = 0$ occurs either when at least one component of $\boldsymbol{u}$ equals 0 or at least $d-1$ components equal 1. The right-hand side further equals

$$Z_n \Big[\int_0^{Ka_n} g_\omega\{\psi_\theta(\boldsymbol{w}x)\} dx + \int_0^{Ka_n} g_\omega\{\psi_{\theta_n}(\boldsymbol{w}x)\} - g_\omega\{\psi_\theta(\boldsymbol{w}x)\} dx \Big] .$$

From the proof of Lemma 4.7 (Treatment of I_4), this is bounded above by

$$Z_n d \int_0^{K\delta} \{1 - \psi_\theta(x)\}^\omega dx + Z_n \sup_{\boldsymbol{w}\in B_{1/K}} \int_0^{x_\Psi/w_{(d)}} |g_\omega\{\psi_{\theta_n}(\boldsymbol{w}x)\} - g_\omega\{\psi_\theta(\boldsymbol{w}x)\}| dx.$$

Now fix an arbitrary $\varepsilon > 0$ and pick a $\delta > 0$ so that $K\delta < x_\Psi$. Then

$$\Pr(\sup_{\boldsymbol{w}\in B_{1/K}} I_4(\boldsymbol{w}) > \varepsilon) \le \Pr(\sup_{\boldsymbol{w}\in B_{1/K}} I_4(\boldsymbol{w}) > \varepsilon, a_n \le \delta) + \Pr(a_n > \delta) .$$

Given that $a_n \to 0$ in probability by Lemma 5.2 (i), it suffices to show that the first term on the right-hand side tends to 0 as $n \to \infty$. Write

$$
\begin{aligned}
&\Pr(\sup_{\boldsymbol{w}\in B_{1/K}} I_4(\boldsymbol{w}) > \varepsilon, a_n \le \delta) \\
&\le \Pr\Big[Z_n d \int_0^{K\delta} \{1-\psi_\theta(x)\}^\omega dx > \frac{\varepsilon}{2}\Big] \\
&\quad + \Pr\Big[Z_n \sup_{\boldsymbol{w}\in B_{1/K}} \int_0^{x_\Psi/w_{(d)}} |g_\omega\{\psi_{\theta_n}(\boldsymbol{w}x)\} - g_\omega\{\psi_\theta(\boldsymbol{w}x)\}| dx > \frac{\varepsilon}{2}\Big] .
\end{aligned}
$$

Given that $Z_n \rightsquigarrow Z = \sup_{\boldsymbol{u}\in[0,1]^d} |\mathbb{C}(\boldsymbol{u})/\tilde{g}_\omega(\boldsymbol{u})|$ as $n \to \infty$ by Theorem 2.12, the Portmanteau lemma implies that the lim sup as $n \to \infty$ of the first term is bounded above by

$$
\Pr\Big[Zd \int_0^{K\delta} \{1-\psi_\theta(x)\}^\omega dx \ge \frac{\varepsilon}{2}\Big] \le \Pr\Big[ZdK\delta \ge \frac{\varepsilon}{2}\Big].
$$

The last probability tends to 0 as $\delta \to 0$. Lemma 5.3 (ii) and the fact that $Z_n \rightsquigarrow Z$ imply that

$$
\lim_{n\to\infty} \Pr\Big[Z_n \sup_{\boldsymbol{w}\in B_{1/K}} \int_0^{x_\Psi/w_{(d)}} |g_\omega\{\psi_{\theta_n}(\boldsymbol{w}x)\} - g_\omega\{\psi_\theta(\boldsymbol{w}x)\}| dx > \frac{\varepsilon}{2}\Big] = 0
$$

which concludes that $\sup_{\boldsymbol{w}\in B_{1/K}} I_4(\boldsymbol{w}) \to 0$ in probability as $n \to \infty$.

Treatment of I_5. We can proceed similarly as when treating I_4. Fix an any $\boldsymbol{w} \in B_{1/K}$ and suppose that $b_n > \delta$ for some $\delta \in (0, x_\Psi)$ arbitrarily close to x_Ψ. Using the arguments from the proof of Lemma 4.7 (treatment of I_5), one has that

$$
\begin{aligned}
I_5(\boldsymbol{w}) &= \int_{b_n/w_{(d)}}^{x_\Psi/w_{(d)}} \left|\frac{\bar{C}_n\{\psi_{\theta_n}(\boldsymbol{w}x)\}}{\tilde{g}_\omega\{\psi_{\theta_n}(\boldsymbol{w}x)\}}\right| \tilde{g}_\omega\{\psi_{\theta_n}(\boldsymbol{w}x)\} dx \\
&\le Z_n \int_{b_n/w_{(d)}}^{x_\Psi/w_{(d)}} \tilde{g}_\omega\{\psi_{\theta_n}(\boldsymbol{w}x)\} dx = Z_n \int_{b_n/w_{(d)}}^{x_\Psi/w_{(d)}} g_\omega\{\psi_{\theta_n}(\boldsymbol{w}x)\} dx,
\end{aligned}
$$

and that the upper bound is bounded above by

$$
Z_n \Big[\int_\delta^{x_\Psi} \{\psi_\theta(x)\}^\omega dx + \sup_{\boldsymbol{w}\in B_{1/K}} \int_0^{x_\Psi/w_{(d)}} |g_\omega\{\psi_{\theta_n}(\boldsymbol{w}x)\} - g_\omega\{\psi_\theta(\boldsymbol{w}x)\}| dx\Big].
$$

The fact that $\sup_{\boldsymbol{w}\in B_{1/K}} I_5(\boldsymbol{w}) \to 0$ in probability as $n \to \infty$ can now be shown using the same arguments as were used in the preceding paragraph to prove that $\sup_{\boldsymbol{w}\in B_{1/K}} I_4(\boldsymbol{w}) \to 0$ in probability as $n \to \infty$, given that $b_n \to x_\Psi$ in probability as $n \to \infty$ by Lemma 5.2 (i) and that $\int_\delta^{x_\Psi} \{\psi_\theta(x)\}^\omega dx \to 0$ as $\delta \to x_\Psi$ by Lemma 4.4. □

Finally, introduce $\check{\mathbb{B}}^{\mathrm{P}}_{n1}$ given, for all $\boldsymbol{w} \in \Delta_d$, by

$$
\check{\mathbb{B}}^{\mathrm{P}}_{n1}(\boldsymbol{w}) = \{\mathrm{E}(Z)\}^{-1} \int_0^{x_\Psi/w_{(d)}} \bar{\mathbb{C}}_n\{\psi_\theta(\boldsymbol{w}x)\} dx .
$$

which by Lemma 5.11 behaves asymptotically as $\bar{\mathbb{B}}^{\mathrm{P}}_{n1}$.

Lemma 5.11. *As $n \to \infty$, $\sup_{\boldsymbol{w}\in B_{1/K}} |\bar{\mathbb{B}}_{n1}^{P}(\boldsymbol{w}) - \check{\mathbb{B}}_{n1}^{P}(\boldsymbol{w})|$ converges to 0 in probability.*

Proof. Introduce the process $\tilde{\mathbb{B}}_{n1}^{\mathrm{P}}$ given, for all $\boldsymbol{w} \in \Delta_d$, by

$$\tilde{\mathbb{B}}_{n1}^{\mathrm{P}}(\boldsymbol{w}) = \{\mathrm{E}(Z)\}^{-1} \int_0^{x_\Psi/w_{(d)}} \frac{\bar{\mathbb{C}}_n\{\psi_{\theta_n}(\boldsymbol{w}x)\}}{\tilde{g}_\omega\{\psi_{\theta_n}(\boldsymbol{w}x)\}} \tilde{g}_\omega\{\psi_\theta(\boldsymbol{w}x)\}dx$$

and observe that, with Z_n as in (4.16),

$$\sup_{\boldsymbol{w}\in B_{1/K}} |\bar{\mathbb{B}}_{n1}^{\mathrm{P}}(\boldsymbol{w}) - \tilde{\mathbb{B}}_{n1}^{\mathrm{P}}(\boldsymbol{w})| \leq \\ Z_n\{\mathrm{E}(Z)\}^{-1} \sup_{\boldsymbol{w}\in B_{1/K}} \int_0^{x_\Psi/w_{(d)}} |\tilde{g}_\omega\{\psi_{\theta_n}(\boldsymbol{w}x)\} - \tilde{g}_\omega\{\psi_\theta(\boldsymbol{w}x)\}|dx$$

From Theorem 2.12, Z_n converges in law to $\sup_{\boldsymbol{u}\in[0,1]^d} |\mathbb{C}(\boldsymbol{u})/\tilde{g}_\omega(\boldsymbol{u})|$ as $n \to \infty$. Furthermore, because $\phi_\theta(0) = \phi_{\theta_n}(0) = x_\Psi$ from Condition 5.3,

$$\sup_{\boldsymbol{w}\in B_{1/K}} \int_0^{x_\Psi/w_{(d)}} |\tilde{g}_\omega\{\psi_{\theta_n}(\boldsymbol{w}x)\} - \tilde{g}_\omega\{\psi_\theta(\boldsymbol{w}x)\}|dx = \\ \sup_{\boldsymbol{w}\in B_{1/K}} \int_0^{x_\Psi/w_{(d)}} |g_\omega\{\psi_{\theta_n}(\boldsymbol{w}x)\} - g_\omega\{\psi_\theta(\boldsymbol{w}x)\}|dx.$$

The expression on the right-hand side tends to zero in probability by Lemma 5.3 (ii). Consequently, $\sup_{\boldsymbol{w}\in B_{1/K}} |\bar{\mathbb{B}}_{n1}^{\mathrm{P}}(\boldsymbol{w}) - \tilde{\mathbb{B}}_{n1}^{\mathrm{P}}(\boldsymbol{w})|$ converges to 0 in probability as $n \to \infty$. Next, recall that the sequence (T_n) with T_n as in (5.32) is tight. Hence, for any $\delta > 0$ there exists $M_\delta > 0$ and $N_\delta > 0$ such that for all $n \geq N_\delta$, $\Pr(T_n > M_\delta) < \delta$. Let $\varepsilon > 0$ be arbitrary. Then

$$\Pr(\sup_{\boldsymbol{w}\in B_{1/K}} |\check{\mathbb{B}}_{n1}^{\mathrm{P}}(\boldsymbol{w}) - \tilde{\mathbb{B}}_{n1}^{\mathrm{P}}(\boldsymbol{w})| > \varepsilon) \leq \delta + \\ \Pr\Big\{\int_0^{x_\Psi/w_{(d)}} \tilde{g}_\omega\left\{\psi_\theta(\boldsymbol{w}x)\right\} dx \sup_{\substack{\boldsymbol{u},\boldsymbol{u}'\in[0,1]^d \\ \|\boldsymbol{u}-\boldsymbol{u}'\|_2 \leq M_\delta/\sqrt{n}}} \Big|\frac{\bar{\mathbb{C}}_n(\boldsymbol{u})}{\tilde{g}_\omega(\boldsymbol{u})} - \frac{\bar{\mathbb{C}}_n(\boldsymbol{u}')}{\tilde{g}_\omega(\boldsymbol{u}')}\Big| > \varepsilon\Big\}.$$

As shown in (4.18), $\int_0^{x_\Psi/w_{(d)}} \tilde{g}_\omega\left\{\psi_\theta(\boldsymbol{w}x)\right\} dx$ is bounded. Because $\delta > 0$ was arbitrary, the conclusion follows from Equation (4.2) of Berghaus et al. (2017). □

Combining the above lemmas,

$$\sup_{\boldsymbol{w}\in B_{1/K}} \|\hat{\mathbb{B}}_n^{\mathrm{P}}(\boldsymbol{w}) - \check{\mathbb{B}}_{n1}^{\mathrm{P}}(\boldsymbol{w}) - \check{\mathbb{B}}_{n2}^{\mathrm{P}}(\boldsymbol{w})\| = o_P(1).$$

Equation (5.12) and the Continuous Mapping Theorem then imply that

$$\hat{\mathbb{B}}_n^{\mathrm{P}} \quad \rightsquigarrow \quad \frac{1}{\mathrm{E}(Z)}\Big[\int_0^{x_\Psi/w_{(d)}} \mathbb{C}\{\psi_\theta(\boldsymbol{w}x)\}dx \;+\; \sum_{j=1}^d \int_0^{x_\Psi/w_j} \dot{C}_j\{\psi_\theta(\boldsymbol{w}x)\}\dot{\psi}_\theta^\top(w_jx)\Theta dx\Big]$$

in $\ell^\infty([0,1]^d)$, as was to be shown. The continuity of the mapping follows from (5.30) and the calculations in the last paragraph of Section 4.2.4. Because for any $j \in \{1,\dots,d\}$, $\dot{C}_j(\boldsymbol{u}) = 0$ if $u_k = 0$ for some $k \neq j$, the limit can be written more succinctly as

$$\frac{1}{\mathrm{E}(Z)} \int_0^{x_\Psi / w_{(d)}} \Big[\mathbb{C}\{\psi_\theta(\boldsymbol{w}x)\} + \sum_{j=1}^d \dot{C}_j\{\psi_\theta(\boldsymbol{w}x)\} \dot{\psi}_\theta^\top(w_j x)\Theta \Big] dx$$

and by change of variable as

$$\frac{1}{\mathrm{E}(Z)} \int_0^1 \Big(\mathbb{C}[\psi_\theta\{-\boldsymbol{w}\log(u)\}] + \sum_{j=1}^d \dot{C}_j[\psi_\theta\{-\boldsymbol{w}\log(u)\}] \dot{\psi}_\theta^\top\{-w_j\log(u)\}\Theta \Big) \frac{du}{u}$$

with the convention that, if $x_\Psi < \infty$, $\dot{\psi}_\theta^\top(x) \equiv \mathbf{0}$ whenever $x \geq x_\Psi$.

Chapter 6

Clustered Archimax model

So far in this thesis, the Archimax model has been advocated as a flexible way to model a group of variables whose asymptotic dependence is driven by a stable tail dependence function; or more precisely a random vector whose dependence structure follows the asymptotic extreme-value regime perturbed by the same distortion. However, as it is the case for rainfall over large territories for example, asymptotic independence between certain variables is likely to be present and this phenomenon cannot be handled by a single Archimax model without limiting the marginal dependence structure to be an (exchangeable) Archimedean copula. Likewise, assuming the same distortion for all variables may not be realistic when the number of variables is large. How to introduce a greater flexibility within the model?

The aim of this chapter is to propose a dependence model in a way that its higher-dimensional margins are Archimax copulas but with possibly different distortions or stable tail dependence functions. To this end, recall that a random vector with stochastic representation (2.11) has an Archimax survival copula; it can thus be seen as a cluster of variables $S_1, \ldots, S_d$ affected by the same random distortion R. Suppose for the moment that the variables $X_1, \ldots, X_d$ can be clustered in a way that each group is a random vector of the form (2.11). This means that each cluster has an Archimax survival copula, with a cluster-specific stdf and distortion variable. The idea pursued here is to introduce dependence between the clusters by making the cluster-specific distortion variables dependent. The advantage of this hierarchical approach is that the entire d-variate copula needs not be constructed explicitly and that within-cluster dependence is Archimax by design.

The clustered Archimax model is introduced formally in Section 6.1; it only concerns the underlying copula and thus has the added flexibility that the margins can be arbitrary. Section 6.2 studies properties of clustered Archimax copulas, specifically how the dependence between the distortions $R_1, \ldots, R_K$ impacts the dependence between the clusters. More importantly, extremal behavior of clustered Archimax copulas is established in the same section. Proofs are reported in Section 6.3, while Section 6.4 formulates a conjecture extending Theorem 6.1.

6.1 Model specification

For a given stdf ℓ and $d \geq 2$, recall first the random vector $\boldsymbol{S} = (S_1, \ldots, S_d)$ with survival function $\bar{G}_d$ of the form (2.10), that is, for all $\boldsymbol{s} \in [0,1]^d$,

$$\bar{G}_d(\boldsymbol{s}) = [\max\{0, 1 - \ell(\boldsymbol{s})\}]^{d-1} . \tag{6.1}$$

Note in particular that the margins of $\boldsymbol{S}$ are Beta; specifically, $S_i \sim B(1, d-1)$ for all $i \in \{1, \ldots, d\}$. Furthermore, let $\mathcal{G} = \{\mathcal{G}_1, \ldots, \mathcal{G}_K\}$ be a partition of $\{1, \ldots, d\}$ into K sets. Because the stochastic representation (2.11) only makes sense in dimensions two and higher, we shall require, throughout this chapter, that $d_k = |\mathcal{G}_k| \geq 2$ for all $k \in \{1, \ldots, K\}$. Hence $K \leq \lfloor d/2 \rfloor$ and of course also $d_1 + \cdots + d_K = d$. Unless stated otherwise, whenever we write $\mathcal{G}_k = \{i_1, \ldots, i_{d_k}\}$ we assume that the indices are ordered, viz. $i_1 < i_2 < \cdots < i_{d_k}$.

As we shall see shortly, a clustered Archimax copula is specified through a partition $\mathcal{G}$ as well as K stdfs and Archimedean generators, respectively. To ease the notation, $\boldsymbol{\ell}$ will denote $(\ell_1, \ldots, \ell_K)$ where for each $k \in \{1, \ldots, K\}$, ℓ_k is a d_k-variate stdf. Similarly, $\boldsymbol{\psi}$ will stand for $(\psi_1, \ldots, \psi_K)$ where for each $k \in \{1, \ldots, K\}$, ψ_k is a d_k-monotone Archimedean generator.

Definition 6.1. *A d-variate copula C is called clustered Archimax copula with cluster partition $\mathcal{G} = \{\mathcal{G}_1, \ldots, \mathcal{G}_K\}$, stdfs $\boldsymbol{\ell}$ and Archimedean generators $\boldsymbol{\psi}$, in notation $C_{\mathcal{G}\boldsymbol{\psi}\boldsymbol{\ell}}$, if it is the survival copula of a random vector $\boldsymbol{X}$ that satisfies the following:*

(i) For each $k \in \{1, \ldots, K\}$ and $i_j \in \mathcal{G}_k = \{i_1, \ldots, i_{d_k}\}$, $X_{i_j} = R_k S_j^{(k)}$ where $\boldsymbol{S}^{(k)} = (S_1^{(k)}, \ldots, S_{d_k}^{(k)})$ has survival function $\bar{G}_k$ and R_k is distributed as the inverse Williamson d_k-transform of ψ_k.

(ii) The random vectors $\boldsymbol{S}^{(1)}, \ldots, \boldsymbol{S}^{(K)}$ are mutually independent.

(iii) The random vector $\boldsymbol{R} = (R_1, \ldots, R_K)$ is independent of $\boldsymbol{S}^{(1)}, \ldots, \boldsymbol{S}^{(K)}$.

As the name suggests, certain multivariate margins of a clustered Archimax copula are Archimax. Specifically, if $\boldsymbol{X}$ is as in Definition 6.1, Theorem 2.9 ensures that for each $k \in \{1, \ldots, K\}$ with $\mathcal{G}_k = \{i_1, \ldots, i_{d_k}\}$, the survival copula of $(X_{i_1}, \ldots, X_{i_{d_k}})$ is the d_k-dimensional Archimax copula $C_{\psi_k \ell_k}$. In particular, in the boundary case when $K = 1$, the entire copula is Archimax.

Before we investigate clustered Archimax copulas in more detail in the next section, we will henceforth assume for simplicity that the partition $\mathcal{G}$ is contiguous. This means that $\mathcal{G}_1 = \{1, \ldots, d_1\}$, $\mathcal{G}_2 = \{d_1 + 1, \ldots, d_1 + d_2\}$ and so on, and leads to no loss of generality. The random vector $\boldsymbol{X}$ in Definition 6.1 is then

$$\left(R_1 S_1^{(1)}, \ldots, R_1 S_{d_1}^{(1)}, \ldots, R_K S_1^{(K)}, \ldots, R_K S_{d_K}^{(K)}\right) . \tag{6.2}$$

Furthermore, from the proof of Theorem 3.3 in Charpentier et al. (2014), the clustered Archimax copula C with cluster partition $\mathcal{G} = \{\mathcal{G}_1, \ldots, \mathcal{G}_K\}$, stdfs $\boldsymbol{\ell}$ and Archimedean generators $\boldsymbol{\psi}$ is the distribution function of

$$\left(\psi_1(R_1 S_1^{(1)}), \ldots, \psi_1(R_1 S_{d_1}^{(1)}), \ldots, \psi_K(R_K S_1^{(K)}), \ldots, \psi_K(R_K S_{d_K}^{(K)})\right). \tag{6.3}$$

6.2 Model properties

In this section, we investigate the extremal behavior of a clustered Archimax copula $C_{\mathcal{G},\boldsymbol{\psi},\boldsymbol{\ell}}$. The main result, Theorem 6.1 below, delineates the conditions under which $C_{\mathcal{G},\boldsymbol{\psi},\boldsymbol{\ell}}$ is in a copula domain of attraction of some extreme-value copula and identifies the latter. Again, without loss of generality, we shall assume that the partition $\mathcal{G}$ is contiguous. Because $C_{\mathcal{G},\boldsymbol{\psi},\boldsymbol{\ell}}$ is also the copula of $1/\boldsymbol{X}$ with $\boldsymbol{X}$ as in (6.2), extremal behavior of $1/\boldsymbol{X}$ will be needed.

Given a contiguous partition $\mathcal{G}$, we will need to introduce the following indexing of components of (random) vectors. Specifically, we shall write $\boldsymbol{X} = (\boldsymbol{X}^{(1)}, \ldots, \boldsymbol{X}^{(K)})$, where for each $k \in \{1, \ldots, K\}$, $\boldsymbol{X}^{(k)} = (X_1^{(k)}, \ldots, X_{d_k}^{(k)})$. Similarly, we shall partition an arbitrary $\boldsymbol{x} \in \mathbb{R}^d$ as $\boldsymbol{x} = (\boldsymbol{x}^{(1)}, \ldots, \boldsymbol{x}^{(K)})$, where for each $k \in \{1, \ldots, K\}$, $\boldsymbol{x}^{(k)} = (x_1^{(k)}, \ldots, x_{d_k}^{(k)})$. Finally, the margins of a d-variate distribution function H will be denoted as $H_1^{(1)}, \ldots, H_{d_1}^{(1)}, \ldots, H_1^{(K)}, \ldots, H_{d_K}^{(K)}$.

The distortion vector $\boldsymbol{R}$ has an effect on both inter- and intra-cluster dependence at extreme levels. Its extreme behavior is important, so it is natural to make the following two assumptions. The first concerns the properties of the margins of $1/\boldsymbol{R}$.

Assumption 6.1. *For a clustered Archimax copula as in Defintion 6.1, assume that $\{1, \ldots, K\}$ is the union of disjoint sets $\mathcal{D}_1$ and $\mathcal{D}_2$, such that*

(i) $k \in \mathcal{D}_1$ if and only if $1/R_k \in \mathcal{M}(\Phi_{\rho_k})$ for some $\rho_k \in (0, 1)$.

(ii) $k \in \mathcal{D}_2$ if and only if there exists an $\epsilon_k > 0$ such that $\mathrm{E}\{1/R_k^{(1+\epsilon_k)}\} < \infty$.

If Assumption 6.1 holds, $k \in \mathcal{D}_1$ means that $1/R_k$ is heavy-tailed and holds if and only if ψ_k satisfies Condition 3.1 with $m_k = 1/\rho_k > 1$. In contrast, $k \in \mathcal{D}_2$ implies that ψ_k satisfies Condition 3.1 with $m_k = 1$ by Proposition 2 in Belzile and Nešlehová (2017). By the same proposition, one then has that $1/X_i^{(k)} \in \mathcal{M}(\Phi_{\rho_k})$ for $k \in \mathcal{D}_1$ and $i \in \{1, \ldots, d_k\}$ and $1/X_i^{(k)} \in \mathcal{M}(\Phi_{\rho_1})$ for $k \in \mathcal{D}_2$ and $i \in \{1, \ldots, d_k\}$. This means that under Assumption 6.1, the respective clustered Archimax copula is in the copula domain of attraction of an extreme-value copula C_0 if and only if $1/\boldsymbol{X}$ is in the maximum domain of attraction of an extreme-value distribution with copula C_0. Such a domain of attraction result requires further assumptions on the extremal behavior of the entire vector $1/\boldsymbol{R}$.

Assumption 6.2. *For a clustered Archimax copula as in Definition 6.1, assume that the reciprocal distortion vector $1/\boldsymbol{R}$ is in the maximum domain of attraction of a multivariate extreme-value distribution with stable tail dependence function $\ell_{1/\boldsymbol{R}}$ given, for $(x_1, \dots, x_K) \in \mathbb{R}_+^K$, by*

$$\ell_{1/\boldsymbol{R}}(x_1, \dots, x_K) = \mathrm{E}[\max_{k=1,\dots,K} \{x_k W_k\}]$$

for some positive random variables $W_1, \dots, W_K$ with unit mean.

Here, we choose the d-norm representation for stable tail dependence functions as discussed in Aulbach et al. (2015). The characterization of (standard) max-stable distributions can be attributed to Pickands (1975), de Haan and Resnick (1977) and Vatan (1985). We are now in position to formulate the main result of this Chapter.

Theorem 6.1. *Let $C_{\mathcal{G},\boldsymbol{\psi},\boldsymbol{\ell}}$ be a clustered Archimax copula with a contiguous partition $\mathcal{G}$ and such that Assumptions 6.1 and 6.2 hold. For $k \in \mathcal{D}_1$, let $b_k = \mathrm{E}\{(1/Z_k)^{\rho_k}\}$, $Z_k \sim B(1, d_k - 1)$. Then $1/\boldsymbol{X} \in \mathcal{M}(H)$, where the univariate margins of H are $H_i^{(k)} = \Phi_{\rho_k}$ for $k \in \mathcal{D}_1$ and $i \in \{1, \dots, d_k\}$ and $H_i^{(k)} = \Phi_1$ for $k \in \mathcal{D}_2$ and $i \in \{1, \dots, d_k\}$. The stable tail dependence function of H is given for all $\boldsymbol{x} \in \mathbb{R}_+^d$ by*

$$\ell_{\mathcal{G},\boldsymbol{\psi},\boldsymbol{\ell}}(\boldsymbol{x}) = \mathrm{E}\left[\max_{k\in\mathcal{D}_1}\left\{\max_{i=1,\dots,d_k}\left(\frac{x_i^{(k)} W_k}{b_k\{S_i^{(k)}\}^{\rho_k}}\right)\right\}\right] + \sum_{k\in\mathcal{D}_2} \ell_k(x_1^{(k)}, \dots, x_{d_k}^{(k)}) . \tag{6.4}$$

Example 6.1 (Clayton Generator)**.** *Using the inverse Williamson d-transform (see Equation (2.6)), one can obtain the distribution of R in the case when ψ_θ is Clayton with parameter θ. When ψ is d-times differentiable, its inverse Williamson d-transform has the density, given, for $r > 0$, by*

$$f_R(r) = (-1)^d \frac{r^{d-1}\psi^{(d)}(r)}{(d-1)!} ;$$

viz. Eq. (2) in McNeil and Nešlehová (2010). In the Clayton case, one has for $r > 0$,

$$f_R(r) = \frac{\theta^d \left\{\prod_{j=0}^{d}(1/\theta + j)\right\}}{(d-1)!}(1+\theta r)^{-1/\theta - d} r^{d-1} .$$

We can see that for $d \geq 2$ and any $\beta < d$,

$$\mathrm{E}(1/R^\beta) = \frac{\theta^d \left\{\prod_{j=0}^{d}(1/\theta + j)\right\}}{(d-1)!} \int_0^\infty \frac{r^{d-1-\beta}}{(1+\theta r)^{1/\theta + d}} dr < \infty .$$

Thus if the k-th cluster has a Clayton distortion, then its components are asymptotically independent from all other clusters since $k \in \mathcal{D}_2$ in Theorem 6.1.

Example 6.2 (Joe generator). *Recall the form of the Joe generator ψ_θ from Table 2.1. Since $1 - \psi_\theta(1/\cdot) \in \mathcal{R}_{-1/\theta}$, $1/R \in \mathcal{M}(\Phi_{1/\theta})$ by Theorem 2 from Larsson and Nešlehová (2011). Therefore if the k-th cluster has a Joe distortion, then it is asymptotically dependent with all other clusters $j \in \mathcal{D}_1$, whose distortions R_j are asymptotically dependent with R_k.*

Inter-cluster asymptotic independence can also be achieved if the distortions are asymptotically independent, as shown in the following corollary.

Corollary 6.1. *If $\{1/R_j : j \in \mathcal{D}_1\}$ are asymptotically independent, then the limiting stdf in* (6.4) *simplifies to*

$$\ell_{\mathcal{G},\boldsymbol{\psi},\boldsymbol{\ell}}(\boldsymbol{x}) = \sum_{k\in\mathcal{D}_1} \ell_k^{\rho_k}((x_1^{(k)})^{1/\rho_k}, \dots, (x_{d_k}^{(k)})^{1/\rho_k}) + \sum_{k\in\mathcal{D}_2} \ell_k(x_1^{(k)} \dots, x_{d_k}^{(k)}) \; .$$

Remark 6.1. *Note that under the hypothesis of Theorem 6.1, the asymptotic behavior of $\{1/R_k : k \in \mathcal{D}_2\}$ has no influence on the form of $\ell_{\mathcal{G},\boldsymbol{\psi},\boldsymbol{\ell}}$.*

The following corollary to Theorem 6.1 compares the inter-cluster stable tail dependence function to that of the reciprocal distortions $(1/R_1, \dots, 1/R_K)$.

Corollary 6.2. *Under the hypothesis of Theorem 6.1, let $\mathcal{I} = (i_1, \dots, i_K)$ be a vector of indices such that $1 \le i_k \le d_k$ for each $k \in \{1, \dots, K\}$. Then, for all $\boldsymbol{x} \in \mathbb{R}_+^K$,*

$$\ell_{1/\boldsymbol{R}}(\boldsymbol{x}) \le \ell_{\mathcal{G},\boldsymbol{\psi},\boldsymbol{\ell}}(\boldsymbol{x}_\mathcal{I}) \; ,$$

where $\boldsymbol{x}_\mathcal{I} = (\boldsymbol{x}_\mathcal{I}^{(1)}, \dots, \boldsymbol{x}_\mathcal{I}^{(K)})$ is defined as follows: For $k \in \{1, \dots, K\}$, $\boldsymbol{x}_\mathcal{I}^{(k)} = (x_{\mathcal{I},1}^{(K)}, \dots, x_{\mathcal{I},d_k}^{(K)})$ where for each $j \in \{1, \dots, d_k\}$, $x_{\mathcal{I},j}^{(k)} = x_k$ if $j = i_k$ and $x_{\mathcal{I},j}^{(k)} = 0$ otherwise.

Remark 6.2. *The first component of* (6.4) *elicits a new method to combine different stdfs in a non-trivial way. Since the second component of* (6.4) *does not reveal any new combination of stdfs, suppose for now that $\mathcal{D}_2 = \emptyset$. For a given $k \in \{1, \dots, K\}$ (and therefore in $\mathcal{D}_1$), setting $x_i^{(l)} = 0$ for all $l \ne k$ and all $i = 1, \dots, d_l$ recovers the marginal stdf of the cluster k. Recall that $b_k = \mathrm{E}\{(1/Z_k)^{\rho_k}\}$ with $Z_k \sim B(1, d_k - 1)$. This marginal stdf is equal to the following for $(x_1^{(k)}, \dots, x_{d_k}^{(k)}) \in \mathbb{R}_+^{d_k}$,*

$$\mathrm{E}\left[\max_{i=1,\dots,d_k}\left(\frac{x_i^{(k)}}{b_k\{S_i^{(k)}\}^{\rho_k}}\right)\right]$$

which itself is equal to $\ell_k^{\rho_k}\left(\{x_1^{(k)}\}^{1/\rho_k}, \dots, \{x_{d_k}^{(k)}\}^{1/\rho_k}\right)$ by Proposition 2.1. In the bivariate case, the form above is a special case of (7) in Engelke et al. (2019). The complete stdf, defined in $\mathbb{R}_+^d$ by

$$\mathrm{E}\left[\max_{k\in\mathcal{D}_1}\left\{\max_{i=1,\dots,d_k}\left(\frac{x_i^{(k)} W_k}{b_k\{S_i^{(k)}\}^{\rho_k}}\right)\right\}\right]$$

essentially mixes the marginal cluster stdfs $\ell_1^{\rho_1}(\{\boldsymbol{x}^{(1)}\}^{1/\rho_1}), \ldots, \ell_K^{\rho_K}(\{\boldsymbol{x}^{(K)}\}^{1/\rho_K})$ *with the limiting stdf of* $(1/R_1, \ldots, 1/R_K)$*. Corollary 6.2 shows that this mixing results in a weaker asymptotic dependence between clusters than that of the reciprocal distortions* $(1/R_1, \ldots, 1/R_K)$*, characterized by* $\ell_{1/\boldsymbol{R}}$*.*

The clustered Archimax model studied in this chapter is related to several other recent articles in the literature. Hierarchical constructions based on Archimax copulas were proposed by Hofert et al. (2018). Specifically, their construction is based on the frailty representation of Archimax copulas, which only holds for completely monotone generators. Hierarchies can be induced via the frailties, the stdf, or both. It would be interesting to establish the attractor of their proposed hierarchical Archimax copula and compare it to that of the clustered Archimax copula. The extremal dependence structure of Liouville copulas is established in Belzile and Nešlehová (2017). The stochastic representation of Liouville copulas is similar to that of Archimax copulas, as they are survival copulas of vectors of the form $R\boldsymbol{D}$, with R a nonnegative random variable and $\boldsymbol{D}$ a Dirichlet random vector. The work presented in this chapter differs from this by replacing the Dirichlet component by a vector $\boldsymbol{S}$ characterized by an stdf and by allowing for multiple distorting random variables $R_1, \ldots, R_K$, thus inducing a hierarchy (or clustering). Finally, Engelke et al. (2019) establish the extremal dependence of bivariate vectors of the form $R \times (W_1, W_2)$ for an extensive combination of asymptotic behaviors of both R and (W_1, W_2). The attractor of the bivariate Archimax copula is in particular obtained as a special case of their Proposition 1 and equation (6), see Sections 2.1 and 4 therein.

6.3 Proofs

This section contains the proofs of the results from the previous section. We begin with auxiliary results in Section 6.3.1; Theorem 6.1 and its Corollaries are proved in Sections 6.3.2 and 6.3.3, respectively.

6.3.1 Auxiliary results

The following proposition is used to prove Theorem 6.1 but is also of independent interest.

Proposition 6.1. *Let* $\boldsymbol{S} = (S_1, \ldots, S_d)$ *be a random vector with joint survival function* $\bar{G}_d$ *as in* (6.1) *for some stdf* ℓ*. Then* $1/\boldsymbol{S}$ *belongs to the maximum domain of attraction of a multivariate extreme-value distribution with unit Fréchet margins and stdf* ℓ*.*

Proof. For the margins, recall that for each $i \in \{1, \ldots, d\}$, $S_i \sim B(1, d-1)$. The survival function of $1/S_i$ is thus given by $\bar{F}_{1/S_i}(s) = 1 - (1 - 1/s)^{d-1}$; it is easily seen that $\bar{F}_{1/S_i} \in \mathcal{R}_{-1}$. Now set $c_n = \{1 - (1 - 1/n)^{1/(d-1)}\}^{-1}$. From Equation 3.13 in Embrechts et al. (1997), for all $s_i \in \mathbb{R}$, it then holds that $\Pr(1/S_i \leq c_n s_i) \to \Phi_1(s_i)$ as $n \to \infty$. Thus

$1/\boldsymbol{S}$ is in the domain of attraction of a multivariate extreme-value distribution with unit Fréchet margins and stdf ℓ if and only if for all $\boldsymbol{s} \in \mathbb{R}_+^d$,

$$\lim_{n\to\infty} n\left\{1 - \Pr(1/S_1 \le c_n s_1, \ldots, 1/S_d \le c_n s_d)\right\} \\ = \lim_{n\to\infty} n\left[1 - \bar{G}_d\big\{1/(c_n s_1), \ldots, 1/(c_n s_d)\big\}\right] = \ell(1/s_1, \ldots, 1/s_d).$$

To show this, fix an arbitrary $\boldsymbol{s} \in \mathbb{R}_+^d$ and observe that because $c_n \to \infty$ as $n \to \infty$,

$$\bar{G}_d\big\{1/(c_n s_1), \ldots, 1/(c_n s_d)\big\} = \big\{1 - (1/c_n)\ell(1/s_1, \ldots, 1/s_d)\big\}^{d-1}$$

for all n sufficiently large. Now note that as $n \to \infty$, n/c_n^k converges to 0 for all $k \in \{2, \ldots, d-1\}$ and to $1/(d-1)$ for $k = 1$. Consequently,

$$\lim_{n\to\infty} n\left[1 - \big\{1 - (1/c_n)\ell(1/s_1, \ldots, 1/s_d)\big\}^{d-1}\right] \\ = \lim_{n\to\infty} \sum_{k=1}^{d-1} \binom{d-1}{k} (-1)^{k+1} \frac{n}{c_n^k} \ell^k(1/s_1, \ldots, 1/s_d) = \ell(1/s_1, \ldots, 1/s_d)$$

as claimed. □

The following lemma determines the normalizing sequences needed for the proof of Theorem 6.1.

Lemma 6.1. *Let $C_{\mathcal{G},\boldsymbol{\psi},\boldsymbol{\ell}}$ be a clustered Archimax copula with a contiguous partition $\mathcal{G}$ and such that Assumptions 6.1 and 6.2 are satisfied. Then the following hold:*

- *(i) For each $k \in \mathcal{D}_1$ and $i \in \{1, \ldots, d_k\}$, $1/(R_k S_i^{(k)}) \in \mathcal{M}(\Phi_{\rho_k})$. Recall that for $k \in \mathcal{D}_1$, $b_k = \mathrm{E}\{(1/Z_k)^{\rho_k}\}$ where $Z_k \sim B(1, d_k - 1)$. Moreover, there exists a sequence of positive constants $\{a_{nk}\}$ such that for all $x > 0$, $n\Pr(1/R_k > a_{nk}x) \to x^{-\rho_k}$ as $n \to \infty$ and $n\Pr(1/(R_k S_i^{(k)}) > a_{nk} b_k^{1/\rho_k} x) \to x^{-\rho_k}$ as $n \to \infty$.*
- *(ii) For each $k \in \mathcal{D}_2$ and $i \in \{1, \ldots, d_k\}$, $1/(R_k S_i^{(k)}) \in \mathcal{M}(\Phi_1)$. Moreover, there exists a sequence of positive constants $\{a_{nk}\}$ such that for all $x > 0$, $n\Pr(1/S_i^{(k)} > a_{nk}x) \to x^{-1}$ as $n \to \infty$ and $n\Pr(1/(R_k S_i^{(k)}) > a_{nk} b_k x) \to x^{-1}$ as $n \to \infty$, where $b_k = \mathrm{E}\{1/R_k\}$.*

Proof. (i) Let $k \in \mathcal{D}_1$ and $i \in \{1, \ldots, d_k\}$. We then have $(1/R_k) \in \mathcal{M}(\Phi_{\rho_k})$ by assumption and $1/S_i^{(k)} \in \mathcal{M}(\Phi_1)$ owing to the fact that $S_i^{(k)} \sim B(1, d-1)$. By Proposition 3.1.1 in Embrechts et al. (1997), there exists a sequence of positive constants $\{a_{nk}\}$ such that for all $x > 0$, $n\Pr(1/R_k > a_{nk}x) \to x^{-\rho_k}$ as $n \to \infty$. Because $\rho_k < 1$, $\mathrm{E}(1/S_i^{(k)})^{\rho_k + \varepsilon} < \infty$ for some ε sufficiently small. Using the lemma of Breiman (1965) and the fact that $b_k = \mathrm{E}\{(1/S_i^{(k)})^{\rho_k}\}$, we then have, for all $x > 0$,

$$\lim_{n\to\infty} n\Pr\Big(\frac{1}{R_k S_i^{(k)}} > a_{nk} b_k^{1/\rho_k} x\Big) = \\ \lim_{n\to\infty} n\Pr\Big(\frac{1}{R_k} > a_{nk} b_k^{1/\rho_k} x\Big) \frac{\Pr\Big(\frac{1}{R_k S_i^{(k)}} > a_{nk} b_k^{1/\rho_k} x\Big)}{\Pr\Big(\frac{1}{R_k} > a_{nk} b_k^{1/\rho_k} x\Big)} = (x b_k^{1/\rho_k})^{-\rho_k} b_k = x^{-\rho_k}\,. \quad (6.5)$$

Indeed, $n\Pr(1/R_k > a_{nk} b_k^{1/\rho_k} x) \to (x b_k^{1/\rho_k})^{-\rho_k}$ as $n \to \infty$ by the choice of normalizing constants $\{a_{nk}\}$. The convergence of the fraction in the above display is due to Breiman's Lemma. Theorem 2.4 implies that since $1/R_k \in \mathcal{M}(\Phi_{\rho_k})$ and $\rho_k \in (0,1)$, $\bar{F}_{1/R_k} \in \mathcal{R}_{-\rho_k}$. We also have that $1/S_i^{(k)}$ and $1/R_k$ are independent, positive, and $\mathrm{E}[\{1/S_i^{(k)}\}^\gamma] < \infty$ for $\gamma \in (\rho_k, 1)$. By Breiman's lemma, $1/(R_k S_i^{(k)}) \in \mathcal{M}(\Phi_{\rho_k})$ and

$$\frac{\Pr\Big(\frac{1}{R_k S_i^{(k)}} > a_{nk} b_k^{1/\rho_k} x\Big)}{\Pr\Big(\frac{1}{R_k} > a_{nk} b_k^{1/\rho_k} x\Big)} \to \mathrm{E}(\{S_i^{(k)}\}^{-\rho_k}) = b_k$$

as $n \to \infty$.

(ii) Let $k \in \mathcal{D}_2$ and $i \in \{1,\ldots,d_k\}$. The proof of the result relies again on Breiman's lemma; see also Proposition 2(b) of Belzile and Nešlehová (2017). Since $1/S_i^{(k)} \in \mathcal{M}(\Phi_1)$, Proposition 3.1.1 in Embrechts et al. (1997) implies that there exist sequences of positive constants $\{a_{nk}\}$ such that for all $x > 0$, $n\Pr(1/S_i^{(k)} > a_{nk} x) \to x^{-1}$ as $n \to \infty$, and this for all $i = 1,\ldots,d_k$. Recall that $b_k = \mathrm{E}(1/R_k)$. Similarly to the proof of part (i), Breiman's lemma then implies that for all $x > 0$,

$$\begin{aligned}
&\lim_{n\to\infty} n\Pr\Big(\frac{1}{R_k S_i^{(k)}} > a_{nk} b_k x\Big) \\
&\quad = \lim_{n\to\infty} n\Pr\Big(\frac{1}{S_i^{(k)}} > a_{nk} b_k x\Big) \frac{\Pr\Big(\frac{1}{R_k S_i^{(k)}} > a_{nk} b_k x\Big)}{\Pr\Big(\frac{1}{S_i^{(k)}} > a_{nk} b_k x\Big)} = (x b_k)^{-1} b_k = x^{-1}\,. \quad (6.6)
\end{aligned}$$

The convergence of the first part of the above is due to the choice of the normalizing constants $\{a_{nk}\}$. For the convergence of the second term, note that $\bar{F}_{1/S_i^{(k)}} \in \mathcal{R}_{-1}$ and by assumption, $\mathrm{E}\{1/R_k^{1+\epsilon_k}\}$ for some $\epsilon_k > 0$. Finally, since $1/S_i^{(k)}$ and $1/R_k$ are independent and positive, Breiman's lemma implies that $1/(R_k S_i^{(k)}) \in \mathcal{M}(\Phi_1)$ and that

$$\frac{\Pr\Big(\frac{1}{R_k S_i^{(k)}} > a_{nk} b_k x\Big)}{\Pr\Big(\frac{1}{S_i^{(k)}} > a_{nk} b_k x\Big)} \to \mathrm{E}\{1/R_k\} = b_k$$

as $n \to \infty$. This completes the proof. □

The lemma below establishes asymptotic independence between clusters in $\mathcal{D}_1$ and clusters in $\mathcal{D}_2$.

Lemma 6.2. *Suppose that $k \in \mathcal{D}_1$, $l \in \mathcal{D}_2$, $i \in \{1,\ldots,d_k\}$ and $j \in \{1,\ldots,d_l\}$. Let $\{a_{nk}\}$ and $\{a_{nj}\}$ be normalizing sequences as in Lemma 6.1. As in Lemma 6.1 (ii), let $b_l = \mathrm{E}\{1/R_l\}$. Then for all $x, y > 0$,*

$$\lim_{n\to\infty} n\Pr\{1/(R_k S_i^{(k)}) > a_{nk} b_k^{1/\rho_k} x, 1/(R_l S_j^{(l)}) > a_{nl} b_l y\} = 0\,.$$

Proof. Fix $x, y > 0$ and recall that $\rho_k \in (0,1)$. The probability of interest can be written as follows

$$\begin{aligned}
&n\Pr\{1/(R_k S_i^{(k)}) > a_{nk} b_k^{1/\rho_k} x, 1/(R_l S_j^{(l)}) > a_{nl} b_l y\} \\
&= \int_{\mathbb{R}_+^2} n\Pr\{1/S_i^{(k)} > a_{nk} b_k^{1/\rho_k} x r_k, 1/S_j^{(l)} > a_{nl} b_l y r_l\} dF_{R_k,R_l}(r_k, r_l) \\
&= \int_{\mathbb{R}_+^2} n\Pr\{1/S_i^{(k)} > a_{nk} b_k^{1/\rho_k} x r_k\} \Pr\{1/S_j^{(l)} > a_{nl} b_l y r_l\} dF_{R_k,R_l}(r_k, r_l) ,
\end{aligned}$$

where the first equality is due to the independence between (R_k, R_l) and $(S_i^{(k)}, S_j^{(l)})$ and the last equality is due to the independence of $S_i^{(k)}$ and $S_j^{(l)}$. Next, consider the integrand as a sequence of functions $\{f_n\}$ defined on $\mathbb{R}_+^2$. Observe that for each $r_k, r_l > 0$,

$$f_n(r_k, r_l) \le g_n(r_k, r_l) ,$$

where $\{g_n\}$ is itself a sequence of functions on $\mathbb{R}_+^2$ defined by

$$g_n(r_k, r_l) = g_n(r_l) = n\Pr\{1/S_j^{(l)} > a_{nl} b_l y r_l\} .$$

From the choice of $\{a_{nl}\}_{n\in\mathbb{N}}$, for all $r_k, r_l > 0$, $\lim\limits_{n\to\infty} g_n(r_k, r_l) = g(r_k, r_l)$, where $g(r_k, r_l) = 1/(b_l y r_l)$. Moreover,

$$\int_{\mathbb{R}_+^2} g(r_k, r_l) dF_{R_k,R_l}(r_k, r_l) = \int_{\mathbb{R}_+^2} \frac{1}{b_l y r_l} dF_{R_k,R_l}(r_k, r_l) = \frac{1}{y} ,$$

and

$$\int_{\mathbb{R}_+^2} g_n(r_k, r_l) dF_{R_k,R_l}(r_k, r_l) = n\Pr\{1/(R_l S_j^{(l)}) > a_{nl} b_l y\} \to \frac{1}{y}$$

as $n \to \infty$. We therefore have a sequence of nonnegative functions $\{g_n\}$ bounding $\{f_n\}$ from above such that

$$\lim_{n\to\infty} \int_{\mathbb{R}_+^2} g_n(r_k, r_l) dF_{R_k,R_l}(r_k, r_l) = \int_{\mathbb{R}_+^2} \lim_{n\to\infty} g_n(r_k, r_l) dF_{R_k,R_l}(r_k, r_l) .$$

Finally, note that

$$f_n(r_k, r_l) = n\Pr\{1/S_i^{(k)} > a_{nk} b_k^{1/\rho_k} x r_k\} \Pr\{1/S_j^{(l)} > a_{nl} b_l y r_l\} \to 0$$

as $n \to \infty$ since

$$\Pr\{1/S_i^{(k)} > a_{nk} b_k^{1/\rho_k} x r_k\} \to 0 \quad \text{and} \quad n\Pr\{1/S_j^{(l)} > a_{nl} b_l y r_l\} \to 1/\{b_l y r_l\}$$

as $n \to \infty$. The desired result then follows by the generalized Lebesgue dominated convergence theorem (see Theorem 1.21 in Kallenberg (2002), for example). □

We now have enough preliminary results in order to prove Theorem 6.1.

6.3.2 Proof of Theorem 6.1

A random vector $(Y_1, \ldots, Y_d)$ is in the maximum domain of attraction of the extreme-value distribution H with Fréchet margins if and only if there exist sequences of positive constants $(a_{ni}) \in (0, \infty)$, $i \in \{1, \ldots, d\}$, so that, for all $(y_1, \ldots, y_d) \in \mathbb{R}_+^d$,

$$\lim_{n\to\infty} n \{1 - \Pr(Y_1 \le a_{n1}y_1, \ldots, Y_d \le a_{nd}y_d)\} = -\ln H(y_1, \ldots, y_d) .$$

This is a multivariate extension of Proposition 3.1.1 in Embrechts et al. (1997), as used in Belzile and Nešlehová (2017). For each $k \in \{1 \ldots, K\}$, set the sequences $\{a_{nk}\}$ as done in Lemma 6.1. Then the fact that the marginals of H are Fréchet follows from the said Lemma. With the normalizing constants now set, the limit of interest is, for any fixed $(x_1^{(1)}, \ldots, x_{d_1}^{(1)}, \ldots, x_1^{(K)}, \ldots, x_{d_K}^{(K)}) \in \mathbb{R}_+^d$,

$$\begin{aligned}\lim_{n\to\infty} n\Big\{1 - \Pr\Big(&1/(R_1S_1^{(1)}) \le a_{n1}b_1^{1/\rho_1}x_1^{(1)}, \ldots, 1/(R_1S_{d_1}^{(1)}) \le a_{n1}b_1^{1/\rho_1}x_{d_1}^{(1)}, \ldots, \\ &1/(R_KS_1^{(K)}) \le a_{nK}b_K^{1/\rho_K}x_1^{(K)}, \ldots, 1/(R_KS_{d_K}^{(K)}) \le a_{nK}b_K^{1/\rho_K}x_{d_K}^{(K)}\Big)\Big\} ,\end{aligned} \tag{6.7}$$

where for $k \in \mathcal{D}_2$, $b_k = \mathrm{E}\{1/R_k\}$ as in Lemma 6.1 (ii) and for ease of notation, $\rho_k = 1$. Let $\mathcal{I} = \{(k, i) : k = 1, \ldots, K, i = 1 \ldots, d_k\}$ and $\mathcal{P}(\mathcal{I})$ denote its power set. Then (6.7) can be rewritten as

$$\lim_{n\to\infty} n \sum_{p\in\mathcal{P}(\mathcal{I})} (-1)^{|p|+1} \Pr\left(\bigcap_{(k,i)\in p} \{1/(R_kS_i^{(k)}) > a_{nk}b_k^{1/\rho_k}x_i^{(k)}\}\right) . \tag{6.8}$$

Let $\mathcal{P}(\mathcal{I})|_{\mathcal{D}_1,\mathcal{D}_2}$ denote the subset of $\mathcal{P}(\mathcal{I})$ such that for all $p \in \mathcal{P}(\mathcal{I})|_{\mathcal{D}_1,\mathcal{D}_2}$, there exists at least one $(k, i) \in p$, and one $(l, j) \in p$ so that $k \in \mathcal{D}_1$ and $l \in \mathcal{D}_2$. Now fix an arbitrary $p \in \mathcal{P}(\mathcal{I})|_{\mathcal{D}_1,\mathcal{D}_2}$ and pick $(k, i), (l, j) \in p$ so that $k \in \mathcal{D}_1$ and $l \in \mathcal{D}_2$. Then for all $\{x_a^{(c)} : (c, a) \in p\} \in \mathbb{R}_+^{|p|}$,

$$\begin{aligned}&n \Pr\left(\bigcap_{(c,a)\in p} \{1/(R_cS_a^{(c)}) > a_{nc}b_c^{1/\rho_c}x_a^{(c)}\}\right) \\ &\qquad \le n \Pr\{1/(R_kS_i^{(k)}) > a_{nk}b_k^{1/\rho_k}x_i^{(k)}, 1/(R_lS_j^{(l)}) > a_{nl}b_l^{1/\rho_l}x_j^{(l)}\} \to 0\end{aligned}$$

as $n \to \infty$ by Lemma 6.2. Thus the summands in (6.8) for which $p \in \mathcal{P}(\mathcal{I})|_{\mathcal{D}_1,\mathcal{D}_2}$ are asymptotically negligible.

Now let $\mathcal{P}(\mathcal{I})|_{\mathcal{D}_1}$ be the subset of $\mathcal{P}(\mathcal{I})$ such that for all $p \in \mathcal{P}(\mathcal{I})|_{\mathcal{D}_1}$, $(c, a) \in p$ implies that $c \in \mathcal{D}_1$. In other words, $\mathcal{P}(\mathcal{I})|_{\mathcal{D}_1}$ contains only sets of indices (c, a) with $c \in \mathcal{D}_1$. Let

$N_1 = \sum_{k\in\mathcal{D}_1} d_k$ and rewrite the summands in (6.8) with $p \in \mathcal{P}(\mathcal{I})|_{\mathcal{D}_1}$ as follows:

$$
\begin{aligned}
&n \sum_{p\in\mathcal{P}(\mathcal{I})|_{\mathcal{D}_1}} (-1)^{|p|+1} \Pr\Big\{ \bigcap_{(c,a)\in p} \{1/(R_c S_a^{(c)}) > a_{nc} b_c^{1/\rho_c} x_a^{(c)}\}\Big\} \\
&\quad = n\Big(1 - \Pr\Big[\bigcap_{k\in\mathcal{D}_1} \{1/(R_k S_1^{(k)}) \le a_{nk} b_k^{1/\rho_k} x_1^{(k)}, \ldots, 1/(R_k S_{d_k}^{(k)}) \le a_{nk} b_k^{1/\rho_k} x_{d_k}^{(k)}\}\Big]\Big) \\
&\quad = \int_{[0,1]^{N_1}} n\Big(1 - \Pr\Big[\bigcap_{k\in\mathcal{D}_1} \{1/R_k \le a_{nk} b_k^{1/\rho_k} x_1^{(k)} s_1^{(k)}, \\
&\qquad\qquad \ldots, 1/R_k \le a_{nk} b_k^{1/\rho_k} x_{d_k}^{(k)} s_{d_k}^{(k)}\}\Big]\Big) dF_{\{\boldsymbol{S}^{(k)}:k\in\mathcal{D}_1\}}(\boldsymbol{s}^{(k)} : k \in \mathcal{D}_1) \\
&\quad = \int_{[0,1]^{N_1}} n\Big(1 - \\
&\qquad\qquad \Pr\Big[\bigcap_{k\in\mathcal{D}_1} \{1/R_k \le a_{nk} b_k^{1/\rho_k} \min_{i\in\{1,\ldots,d_k\}} \{x_i^{(k)} s_i^{(k)}\}\}\Big]\Big) dF_{\{\boldsymbol{S}^{(k)}:k\in\mathcal{D}_1\}}(\boldsymbol{s}^{(k)} : k \in \mathcal{D}_1) \,.
\end{aligned}
$$

Now consider the integrand as a sequence of functions $\{f_n\}$ defined on $[0,1]^{N_1}$ and observe that for each $n \in \mathbb{N}$, $0 \le f_n \le g_n$, where g_n is given, for each $(\boldsymbol{s}^{(k)} : k \in \mathcal{D}_1) \in [0,1]^{N_1}$, by

$$
g_n(\{\boldsymbol{s}^{(k)} : k \in \mathcal{D}_1\}) = n \sum_{k\in\mathcal{D}_1} \sum_{i=1}^{d_k} \Pr\Big(1/R_k > a_{nk} b_k^{1/\rho_k} x_i^{(k)} s_i^{(k)}\Big)
$$

Clearly,

$$
g_n(\boldsymbol{s}^{(k)} : k \in \mathcal{D}_1) \to g(\boldsymbol{s}^{(k)} : k \in \mathcal{D}_1)
$$

as $n \to \infty$ where

$$
g(\boldsymbol{s}^{(k)} : k \in \mathcal{D}_1) = \sum_{k\in\mathcal{D}_1} \sum_{i=1}^{d_k} \frac{1}{b_k \{x_i^{(k)} s_i^{(k)}\}^{\rho_k}}
$$

with

$$
\int_{[0,1]^{N_1}} g(\boldsymbol{s}^{(k)} : k \in \mathcal{D}_1) dF_{\{\boldsymbol{S}^{(k)}:k\in\mathcal{D}_1\}}(\boldsymbol{s}^{(k)} : k \in \mathcal{D}_1) = \sum_{k\in\mathcal{D}_1} \sum_{i=1}^{d_k} (x_i^{(k)})^{-\rho_k} \,.
$$

Moreover,

$$
\begin{aligned}
&\int_{[0,1]^{N_1}} g_n(\boldsymbol{s}^{(k)} : k \in \mathcal{D}_1) dF_{\{\boldsymbol{S}^{(k)}:k\in\mathcal{D}_1\}}(\boldsymbol{s}^{(k)} : k \in \mathcal{D}_1) \\
&\qquad = n\left\{ \sum_{k\in\mathcal{D}_1} \sum_{i=1}^{d_k} \Pr\Big(1/(R_k S_i^{(k)}) > a_{nk} b_k^{1/\rho_k} x_i^{(k)}\Big) \right\} \to \sum_{k\in\mathcal{D}_1} \sum_{i=1}^{d_k} (x_i^{(k)})^{-\rho_k}
\end{aligned}
$$

as $n \to \infty$. Therefore, we have a sequence of majorants $\{g_n\}$ such that $\lim_{n\to\infty} \int g_n = \int \lim_{n\to\infty} g_n$. Now recall that the vector of distortions $1/\boldsymbol{R}$ has a limiting stdf $\ell_{1/\boldsymbol{R}}$ defined in terms of the positive, unit-mean variables $W_1, \ldots, W_K$ in Assumption 6.2. Therefore, $f_n \to f$ point-wise, where for all $(\boldsymbol{s}^{(k)} : k \in \mathcal{D}_1) \in [0,1]^{N_1}$,

$$
f(\boldsymbol{s}^{(k)} : k \in \mathcal{D}_1) = \mathrm{E}\Big[\max_{k\in\mathcal{D}_1}\Big\{ \frac{W_k}{(b_k^{1/\rho_k} \min_{i=1,\ldots,d_k} \{x_i^{(k)} s_i^{(k)}\})^{\rho_k}} \Big\}\Big] \,.
$$

Now, integrating over the $(\{\boldsymbol{s}^{(k)} : k \in \mathcal{D}_1\})$ yields the following:

$$\int_{[0,1]^{N_1}} f(\boldsymbol{s}^{(k)} : k \in \mathcal{D}_1) dF_{\{\boldsymbol{S}^{(k)}:k\in\mathcal{D}_1\}}(\boldsymbol{s}^{(k)} : k \in \mathcal{D}_1)$$
$$= \mathrm{E}\Big[\max_{k\in\mathcal{D}_1}\Big\{\max_{i=1,\dots,d_k}\Big(\frac{W_k}{b_k\{S_i^{(k)}x_i^{(k)}\}^{\rho_k}}\Big)\Big\}\Big] .$$

Using the generalized Lebesgue dominated convergence theorem, we can thus conclude that for all $(\boldsymbol{x}^{(k)} : k \in \mathcal{D}_1) \in \mathbb{R}_+^{N_1}$,

$$\lim_{n\to\infty} n \sum_{p\in\mathcal{P}(\mathcal{I})|_{\mathcal{D}_1}} (-1)^{|p|+1} \Pr\Big(\bigcap_{a^{(c)}\in p} \{1/(R_c S_a^{(c)}) > a_{nb} b_c^{\rho_c} x_a^{(c)}\}\Big)$$
$$= \mathrm{E}\Big[\max_{k\in\mathcal{D}_1}\Big\{\max_{i=1,\dots,d_k}\Big(\frac{W_k}{b_k\{S_i^{(k)}x_i^{(k)}\}^{\rho_k}}\Big)\Big\}\Big] .$$

Analogously to $\mathcal{P}(\mathcal{I})|_{\mathcal{D}_1}$, let $\mathcal{P}(\mathcal{I})|_{\mathcal{D}_2}$ contain only sets of indices (c, a) with $c \in \mathcal{D}_2$. Let $K_2 = |\mathcal{D}_2|$ and $N_2 = \sum_{k\in\mathcal{D}_2} d_k$, and recall that $\rho_k = 1$ for $k \in \mathcal{D}_2$. Next, $(\{\boldsymbol{x}_i^{(k)} : k \in \mathcal{D}_2\}) \in \mathbb{R}_+^{N_2}$ and rewrite the summands of (6.8) with $p \in \mathcal{P}(\mathcal{I})|_{\mathcal{D}_2}$ as follows:

$$\begin{aligned}
& n \sum_{p\in\mathcal{P}(\mathcal{I})|_{\mathcal{D}_2}} (-1)^{|p|+1} \Pr\Big(\bigcap_{(c,a)\in p} \{1/(R_c S_a^{(c)}) > a_{nc} b_c^{1/\rho_c} x_a^{(c)}\}\Big) \\
&= n\Big[1 - \Pr\Big(\bigcap_{k\in\mathcal{D}_2} \{1/(R_k S_1^{(k)}) \le a_{nk} b_k x_1^{(k)}, \dots, 1/(R_k S_{d_k}^{(k)}) \le a_{nk} b_k x_{d_k}^{(k)}\}\Big)\Big] \\
&= \int_{\mathbb{R}_+^{|\mathcal{D}_1|}} n\Big[1 - \Pr\Big(\bigcap_{k\in\mathcal{D}_2} \{1/S_1^{(k)} \le a_{nk} b_k x_1^{(k)} r_k, \\
&\qquad \dots, 1/S_{d_k}^{(k)} \le a_{nk} b_k x_{d_k}^{(k)} r_k\}\Big)\Big] dF_{\{R_k:k\in\mathcal{D}_2\}}(r_k : k \in \mathcal{D}_2) .
\end{aligned}$$

Now consider the integrand as a sequence of functions $\{f_n\}$ defined on $\mathbb{R}_+^{K_2}$ and observe that for each $n \in \mathbb{N}$, $0 \le f_n \le g_n$, where g_n is given, for all $(r_k : k \in \mathcal{D}_1) \in \mathbb{R}_+^{K_2}$, by

$$g_n(r_k : k \in \mathcal{D}_2) = n\Big\{\sum_{k\in\mathcal{D}_2} \sum_{i=1}^{d_k} \Pr\big(1/S_i^{(k)} > a_{nk} b_k x_i^{(k)} r_k\big)\Big\}$$

Clearly, for all $(r_k : k \in \mathcal{D}_2) \in \mathbb{R}_+^{K_2}$ and as $n \to \infty$,

$$g_n(r_k : k \in \mathcal{D}_2) \to g(r_k : k \in \mathcal{D}_2) = \sum_{k\in\mathcal{D}_2} \sum_{i=1}^{d_k} \frac{1}{b_k x_i^{(k)} r_k} .$$

Furthermore,

$$\int_{\mathbb{R}_+^{m_2}} g(r_k : k \in \mathcal{D}_2) dF_{\{R_k:k\in\mathcal{D}_2\}}(r_k : k \in \mathcal{D}_2) = \sum_{k\in\mathcal{D}_2} \sum_{i=1}^{d_k} \frac{1}{x_i^{(k)}}$$

and as $n \to \infty$,

$$\int_{\mathbb{R}_+^{m_2}} g_n(r_k : k \in \mathcal{D}_2) dF_{\{R_k : k \in \mathcal{D}_2\}}(r_k : k \in \mathcal{D}_2)$$
$$= n\Big\{\sum_{k\in\mathcal{D}_2}\sum_{i=1}^{d_k} \Pr\big(1/(R_k S_i^{(k)}) > a_{nk} b_k x_i^{(k)}\big)\Big\} \to \sum_{k\in\mathcal{D}_1}\sum_{i=1}^{d_k} \frac{1}{x_i^{(k)}} \,.$$

Analogously to the treatment of $\mathcal{P}(\mathcal{I})|_{\mathcal{D}_1}$ we have a sequence of majorants $\{g_n\}$ such that $\lim_{n\to\infty} \int g_n = \int \lim_{n\to\infty} g_n$. It remains to determine the limit of the sequence of functions $\{f_n\}$ defined for all $(r_k : k \in \mathcal{D}_2) \in \mathbb{R}_+^{K_2}$ by

$$n\Big[1 - \Pr\Big(\bigcap_{k\in\mathcal{D}_2} \{1/S_1^{(k)} > a_{nk} b_k x_1^{(k)} r_k, \ldots, 1/S_{d_k}^{(k)} > a_{nk} b_k x_{d_k}^{(k)} r_k\}\Big)\Big]$$

By assumption, $\boldsymbol{S}^{(k)} = (S_1^{(k)}, \ldots, S_{d_k}^{(k)})$ and $\boldsymbol{S}^{(l)} = (S_1^{(l)}, \ldots, S_{d_l}^{(l)})$ are independent if $k \neq l$ and are therefore asymptotically independent as well. Using Proposition 6.1 and the fact that $1/S_i^{(k)} \in \mathcal{M}(\Phi_1)$ for all $k \in \{1, \ldots, K\}$, $i \in \{1, \ldots, d_k\}$ one has that $f_n \to f$ point-wise, where for all $(r_k : k \in \mathcal{D}_1) \in \mathbb{R}_+^{N_2}$,

$$\begin{aligned} f(r_k : k \in \mathcal{D}_1) &= \sum_{k\in\mathcal{D}_2} \ell_k\left(\{b_k x_1^{(k)} r_k\}^{-1}, \ldots, \{b_k x_{d_k}^{(k)} r_k\}^{-1}\right) \\ &= \sum_{k\in\mathcal{D}_2} (b_k r_k)^{-1} \ell_k\left(\{x_1^{(k)}\}^{-1}, \ldots, \{x_{d_k}^{(k)}\}^{-1}\right) \,. \end{aligned}$$

Integrating the limit f yields

$$\int_{\mathbb{R}_+^{m_2}} f(r_k : k \in \mathcal{D}_2) dF_{\{R_k : k \in \mathcal{D}_2\}}(r_k : k \in \mathcal{D}_2) = \sum_{k\in\mathcal{D}_2} \ell_k\left(\{x_1^{(k)}\}^{-1}, \ldots, \{x_{d_k}^{(k)}\}^{-1}\right) \,.$$

Thus for all $(x_1^{(1)}, \ldots, x_{d_1}^{(1)}, \ldots, x_1^{(K)}, \ldots, x_{d_K}^{(K)}) \in \mathbb{R}_+^d$, the limit (6.7) is equal to

$$\mathrm{E}\Big[\max_{k\in\mathcal{D}_1}\Big\{\max_{i=1,\ldots,d_k}\Big(\frac{W_k}{b_k\{S_i^{(k)} x_i^{(k)}\}^{\rho_k}}\Big)\Big\}\Big] + \sum_{k\in\mathcal{D}_2} \ell_k\left(\{x_1^{(k)}\}^{-1}, \ldots, \{x_{d_k}^{(k)}\}^{-1}\right) \,.$$

Recalling that $1/(R_k S_i^{(k)}) \in \mathcal{M}(\Phi_{\rho_k})$, one obtains (6.4) by plugging in the appropriate Fréchet margins. $\square$

6.3.3 Proofs of Corollaries 6.2 and 6.1

Proof of Corollary 6.1. Let $K_1 = |\mathcal{D}_1|$ and recall that $\mathrm{E}[\max_{k\in\mathcal{D}_1} y_k W_k]$ is the limiting stdf of $\{1/R_k : k \in \mathcal{D}_1\}$, defined for all $(y_1, \ldots, y_{K_1}) \in \mathbb{R}_+^{K_1}$. Letting $\{W_k : k \in \mathcal{D}_1\}$ be a (uniformly) random permutation of $(K_1, 0, \ldots, 0)$ yields the independence stdf $\mathrm{E}[\max_{k\in\mathcal{D}_1} y_k W_k] = y_1 + \ldots + y_{m_1}$. Due to the fact that the W_k are independent of all $S_i^{(l)}$, plugging this into (6.4) yields, for all $\boldsymbol{x} \in \mathbb{R}_+^d$,

$$\ell_{\mathcal{G},\boldsymbol{\psi},\boldsymbol{\ell}}(\boldsymbol{x}) = \sum_{k\in\mathcal{D}_1} \mathrm{E}\Big[\max_{i=1,\ldots,d_k}\Big(\frac{x_i^{(k)}}{b_k\{S_i^{(k)}\}^{\rho_k}}\Big)\Big] + \sum_{k\in\mathcal{D}_2} \ell_k(x_1^{(k)}, \ldots, x_{d_k}^{(k)}) \,.$$

By Proposition 2.1, for each $k \in \mathcal{D}_1$,

$$\mathrm{E}\Big[\max_{i=1,\ldots,d_k}\Big(\frac{x_i^{(k)}}{b_k\{S_i^{(k)}\}^{\rho_k}}\Big)\Big] = \ell_k^{\rho_k}\big(\{x_1^{(k)}\}^{1/\rho_k},\ldots,\{x_{d_k}^{(k)}\}^{1/\rho_k}\big)\ .$$

This completes the proof. □

Proof of Corollary 6.2. Fix an arbitrary $\boldsymbol{x} \in \mathbb{R}_+^K$ and observe first that by Assumption 6.2 and the fact that the variables W_k have unit mean,

$$\begin{aligned}\ell_{1/\boldsymbol{R}}(\boldsymbol{x}) = \mathrm{E}\big[\max_{k=1,\ldots,K}\{x_k W_k\}\big] &\le \mathrm{E}\big[\max_{k\in\mathcal{D}_1}\{x_k W_k\} + \sum_{k\in\mathcal{D}_2} x_k W_k\big] \\ &= \mathrm{E}\big[\max_{k\in\mathcal{D}_1}\{x_k W_k\}\big] + \sum_{k\in\mathcal{D}_2} x_k. \quad (6.9)\end{aligned}$$

Next, note that $\ell_{\mathcal{G},\boldsymbol{\psi},\boldsymbol{\ell}}(\boldsymbol{x}_{\mathcal{I}}) = \mathcal{A}(\boldsymbol{x}) + \mathcal{B}(\boldsymbol{x})$, where

$$\mathcal{A}(\boldsymbol{x}) = \sum_{k\in\mathcal{D}_2} \ell_k\big(x_{\mathcal{I},1}^{(k)},\ldots,x_{\mathcal{I},d_k}^{(k)}\big) = \sum_{k\in\mathcal{D}_2} x_k$$

and

$$\mathcal{B}(\boldsymbol{x}) = \mathrm{E}\Big[\max_{k\in\mathcal{D}_1}\Big\{\max_{j=1,\ldots,d_k}\frac{x_{\mathcal{I},j}^{(k)} W_k}{b_k(S_j^{(k)})^{\rho_k}}\Big\}\Big] = \mathrm{E}\Big\{\max_{k\in\mathcal{D}_1}\frac{x_k W_k}{b_k(S_j^{(k)})^{\rho_k}}\Big\}.$$

Because for each $k \in \mathcal{D}_1$, $b_k = \mathrm{E}\{(1/S_j^{(k)})^{\rho_k}\}$, we have that for any $\boldsymbol{w} \in \mathbb{R}_+^K$ and $k \in \mathcal{D}_1$,

$$\mathrm{E}\Big\{\max_{k\in\mathcal{D}_1}\frac{x_k w_k}{b_k(S_j^{(k)})^{\rho_k}}\Big\} \ge \mathrm{E}\Big\{\frac{x_k w_k}{b_k(S_j^{(k)})^{\rho_k}}\Big\} = x_k w_k\ ,$$

so that

$$\mathrm{E}\Big\{\max_{k\in\mathcal{D}_1}\frac{x_k w_k}{b_k(S_j^{(k)})^{\rho_k}}\Big\} \ge \max_{k\in\mathcal{D}_1}\{x_k w_k\}.$$

This implies that

$$\mathcal{B}(\boldsymbol{x}) \ge \mathrm{E}\big[\max_{k\in\mathcal{D}_1}\{x_k W_k\}\big]$$

which together with (6.9) yields the desired result. □

6.4 Conjectured extension of Theorem 6.1

As it is stated, Theorem 6.1 does not account for the boundary case when $1/R_k \in \mathcal{M}(\Phi_1)$, which can occur. It would thus be desirable to replace Assumption 6.1 of Theorem 6.1 by the following requirement.

Assumption 6.3. *For a clustered Archimax copula as in Definition 6.1, assume that* $\{1,\ldots,K\}$ *is the union of disjoint sets* $\mathcal{D}_1$*,* $\mathcal{D}_2$*, and* $\mathcal{D}_3$ *such that*

(i) $k \in \mathcal{D}_1$ *if and only if* $1/R_k \in \mathcal{M}(\Phi_{\rho_k})$ *for some* $\rho_k \in (0,1)$*.*

(ii) $k \in \mathcal{D}_2$ *if and only if there exists an* $\epsilon_k > 0$ *such that* $\mathrm{E}\{1/R_k^{(1+\epsilon_k)}\} < \infty$.

(iii) $k \in \mathcal{D}_3$ *if and only if* $1/R_k \in \mathcal{M}(\Phi_1)$ *and* $\mathrm{E}\{1/R_k\} = \infty$.

We conjecture that the variables whose distortions are in $\mathcal{D}_3$ have the same asymptotic behavior as those whose distortions are in $\mathcal{D}_2$. More precisely, we surmise that the following statement holds.

Conjecture 6.1. *Let* $C_{\mathcal{G},\boldsymbol{\psi},\boldsymbol{\ell}}$ *be a clustered Archimax copula with a contiguous partition* $\mathcal{G}$ *and such that Assumptions 6.3 and 6.2 hold. For* $k \in \mathcal{D}_1$*, let* $b_k = \mathrm{E}\{(1/Z_k)^{\rho_k}\}$, $Z_k \sim B(1, d_k - 1)$*. Then* $1/\boldsymbol{X} \in \mathcal{M}(H)$*, where the univariate margins of* H *are* $H_i^{(k)} = \Phi_{\rho_k}$ *for* $k \in \mathcal{D}_1$ *and* $i \in \{1,\dots,d_k\}$ *and* $H_i^{(k)} = \Phi_1$ *for* $k \in \mathcal{D}_2 \cup \mathcal{D}_3$ *and* $i \in \{1,\dots,d_k\}$*. The stable tail dependence function of* H *is given for all* $\boldsymbol{x} \in \mathbb{R}_+^d$ *by*

$$\ell_{\mathcal{G},\boldsymbol{\psi},\boldsymbol{\ell}}(\boldsymbol{x}) = \mathrm{E}\left[\max_{k \in \mathcal{D}_1}\left\{\max_{i=1,\dots,d_k}\left(\frac{x_i^{(k)} W_k}{b_k\{S_i^{(k)}\}^{\rho_k}}\right)\right\}\right] + \sum_{k \in \mathcal{D}_2 \cup \mathcal{D}_3} \ell_k(x_1^{(k)},\dots,x_{d_k}^{(k)}) . \tag{6.10}$$

One part of Conjecture 6.1 is clear, namely that $H_i^{(k)} = \Phi_1$ for $k \in \mathcal{D}_3$. Indeed, for any such k, the Corollary to Theorem 3 in Embrechts and Goldie (1980) implies that $1/(R_k S_i^{(k)}) \in \mathcal{M}(\Phi_1)$. So one can again find a sequence $\{a_{nk}\}$ of positive constants ensuring that for all $x \in \mathbb{R}_+$, $n\Pr(1/(R_k S_i^{(k)}) > a_{nk}x) \to 1/x$ as $n \to \infty$. The main difficulty in establishing the validity of Conjecture 6.1 that arises is the fact that, for $k \in \mathcal{D}_3$ and $i \in \{1,\dots,d_k\}$, the relation between the above normalizing sequence $\{a_{nk}\}$ and the normalizing sequences for $1/R_k$, $1/S_i^{(k)}$ is unclear.

In order to prove the conjectured result, it suffices to prove the following three sister lemmas. The first two, analogous to Lemma 6.2, are proved below. The third, conjecturing asymptotic independence between different clusters in $\mathcal{D}_3$, is the missing result that if established would prove that Conjecture 6.1 is indeed true.

Lemma 6.3. *Under the hypothesis of Conjecture 6.1, suppose that* $k \in \mathcal{D}_1$, $l \in \mathcal{D}_3$, $i \in \{1,\dots,d_k\}$ *and* $j \in \{1,\dots,d_l\}$*. Let* $\{a_{nk}\}$ *be a sequence of positive constants such that for all* $x > 0$, $n\Pr(1/R_k > a_{nk}x) \to x^{-\rho_k}$ *as* $n \to \infty$ *and* $n\Pr(1/(R_k S_i^{(k)}) > a_{nk} b_k^{1/\rho_k} x) \to x^{-\rho_k}$ *as* $n \to \infty$*. Furthermore, let* $\{a_{nl}\}$ *be a sequence of positive constants so that for all* $x > 0$, $n\Pr(1/(R_l S_i^{(l)}) > a_{nl}x) \to 1/x$ *as* $n \to \infty$*. Then for all* $x, y \in \mathbb{R}_+$,

$$\lim_{n\to\infty} n\Pr\{1/(R_k S_i^{(k)}) > a_{nk} b_k^{1/\rho_k} x, 1/(R_l S_j^{(l)}) > a_{nl} y\} = 0 .$$

Proof. The proof is quite similar to the one of Lemma 6.2. Observe first that the assumed sequences $\{a_{nk}\}$ and $\{a_{nl}\}$ indeed exist, by Lemma 6.1 and the discussion in the paragraph following Conjecture 6.1. Fix some arbitrary $x, y > 0$ and recall that $\rho_k \in (0,1)$. The probability of interest can be written as follows

$$\begin{aligned} &n\Pr\{1/(R_k S_i^{(k)}) > a_{nk} b_k^{1/\rho_k} x, 1/(R_l S_j^{(l)}) > a_{nl} b_l y\} \\ &\qquad = \int_{(0,1)^2} n\Pr\{1/R_k > a_{nk} b_k^{1/\rho_k} x s_i^{(k)}, 1/R_l > a_{nl} y s_j^{(l)}\} dF_{S_i^{(k)},S_j^{(l)}}(s_i^{(k)}, s_j^{(l)}) \end{aligned}$$

Consider the integrand as a function f_n defined on $(0,1)^2$ and note that for all $n \in \mathbb{N}$, $0 \le f_n \le g_n$, where g_n is given, for all $(s_i^{(k)}, s_j^{(l)}) \in (0,1)^2$ by

$$g_n(s_i^{(k)}, s_j^{(l)}) = g_n(s_i^{(k)}) = n \Pr(1/R_k > a_{nk} b_k^{1/\rho_k} x s_i^{(k)}) \ .$$

As in the proof of Lemma 6.2, for all $(s_i^{(k)}, s_j^{(l)}) \in (0,1)^2$,

$$\lim_{n\to\infty} g_n(s_i^{(k)}, s_j^{(l)}) = g(s_i^{(k)}, s_j^{(l)}) = 1/\{b_k (x s_i^{(k)})^{\rho_k}\} \ .$$

Moreover,

$$\int_{(0,1)^2} g(s_i^{(k)}, s_j^{(l)}) dF_{S_i^{(k)}, S_j^{(l)}}(s_i^{(k)}, s_j^{(l)}) = \frac{1}{x^{\rho_k}}$$

and

$$\int_{(0,1)^2} g_n(s_i^{(k)}, s_j^{(l)}) dF_{S_i^{(k)}, S_j^{(l)}}(s_i^{(k)}, s_j^{(l)}) = n \Pr\{1/(R_k S_i^{(k)}) > a_{nk} b_k^{1/\rho_k} x\} \to \frac{1}{x^{\rho_k}}$$

as $n \to \infty$. We therefore have a sequence of functions $\{g_n\}$ bounding $\{f_n\}$ from above such that

$$\lim_{n\to\infty} \int_{(0,1)^2} g_n(s_i^{(k)}, s_j^{(l)}) dF_{S_i^{(k)}, S_j^{(l)}}(s_i^{(k)}, s_j^{(l)}) = \int_{(0,1)^2} \lim_{n\to\infty} g_n(s_i^{(k)}, s_j^{(l)}) dF_{S_i^{(k)}, S_j^{(l)}}(s_i^{(k)}, s_j^{(l)}) \ .$$

Finally, note that

$$f_n(s_i^{(k)}, s_j^{(l)}) = n \Pr\{1/R_k > a_{nk} b_k^{1/\rho_k} x s_i^{(k)}, 1/R_l > a_{nl} b_l^{1/\rho_l} y s_j^{(l)}\} \to 0$$

as $n \to \infty$ since $n \Pr\{1/R_k > a_{nk} b_k^{1/\rho_k} x s_i^{(k)}\} \to \{b_k^{1/\rho_k} x s_i^{(k)}\}^{-\rho_k}$ and $\Pr\{1/R_l > a_{nl} b_l^{1/\rho_l} y s_j^{(l)}\}$ $\to 0$ as $n \to \infty$. The desired result now follows by the generalized Lebesgue dominated convergence theorem. $\square$

Lemma 6.4. *Under the hypothesis of Conjecture 6.1, suppose that $k \in \mathcal{D}_2$, $l \in \mathcal{D}_3$, $i \in \{1, \ldots, d_k\}$ and $j \in \{1, \ldots, d_l\}$. Let $\{a_{nk}\}$ such that for all $x > 0$, $n \Pr(1/S_i^{(k)} > a_{nk} x) \to x^{-1}$ as $n \to \infty$ and $n \Pr(1/(R_k S_i^{(k)}) > a_{nk} b_k x) \to x^{-1}$ as $n \to \infty$. Furthermore, let $\{a_{nl}\}$ be a sequence of positive constants so that for all $x > 0$, $n \Pr(1/(R_l S_i^{(l)}) > a_{nl} x) \to x^{-1}$ as $n \to \infty$. Then for all $x, y \in \mathbb{R}_+$,*

$$\lim_{n\to\infty} n \Pr\{1/(R_k S_i^{(k)}) > a_{nk} b_k^{1/\rho_k} x, 1/(R_l S_j^{(l)}) > a_{nl} y\} = 0 \ .$$

Proof. This proof is almost exactly the same as the proof of Lemma 6.2. Again, the existence of the norming constants $\{a_{nk}\}$ and $\{a_{nl}\}$ follows from Lemma 6.1 and the discussion in the paragraph following Conjecture 6.1. Fix some arbitrary $x, y > 0$. We are interested in the limit as $n \to \infty$ of

$$\begin{aligned} n \Pr\{1/(R_k S_i^{(k)}) &> a_{nk} b_k x, 1/(R_l S_j^{(l)}) > a_{nl} y\} \\ &= \int_{\mathbb{R}_+^2} n \Pr\{1/S_i^{(k)} > a_{nk} b_k x r_k\} \Pr\{1/S_j^{(l)} > a_{nl} y r_l\} dF_{R_k, R_l}(r_k, r_l) \ . \end{aligned}$$

Consider the integrand as a function f_n defined on $\mathbb{R}_+^2$. Observe that for each $n \in \mathbb{N}$, $0 \le f_n \le g_n$ where for all $(r_k, r_l) \in \mathbb{R}_+^2$,

$$g_n(r_k, r_l) = g_n(r_k) = n \Pr\{1/S_i^{(k)} > a_{nk} b_k x r_k\} \; .$$

From the choice of $\{a_{nk}\}$, for all $(r_k, r_l) \in \mathbb{R}_+^2$,

$$\lim_{n \to \infty} g_n(r_k, r_l) = g(r_k, r_l) = 1/(b_k x r_k) \; .$$

Moreover, since $b_k = \mathrm{E}(1/R_k)$,

$$\int_{\mathbb{R}_+^2} g(r_k, r_l) dF_{R_k, R_l}(r_k, r_l) = \int_{\mathbb{R}_+^2} \frac{1}{b_k x r_k} dF_{R_k, R_l}(r_k, r_l) = \frac{1}{x} \; .$$

and

$$\int_{\mathbb{R}_+^2} g_n(r_k, r_l) dF_{R_k, R_l}(r_k, r_l) = n \Pr\{1/(R_k S_i^{(k)}) > a_{nk} b_k x\} \to \frac{1}{x}$$

as $n \to \infty$. We therefore have a sequence of functions $\{g_n\}$ bounding $\{f_n\}$ from above such that

$$\lim_{n \to \infty} \int_{\mathbb{R}_+^2} g_n(r_k, r_l) dF_{R_k, R_l}(r_k, r_l) = \int_{\mathbb{R}_+^2} \lim_{n \to \infty} g_n(r_k, r_l) dF_{R_k, R_l}(r_k, r_l) \; .$$

Finally, note that

$$f_n(r_k, r_l) = n \Pr\{1/S_i^{(k)} > a_{nk} b_k x r_k\} \Pr\{1/S_j^{(l)} > a_{nl} y r_l\} \to 0$$

as $n \to \infty$ since

$$n \Pr\{1/S_i^{(k)} > a_{nk} b_k x r_k\} \to 1/\{b_k x r_k\} \quad \text{and} \quad \Pr\{1/S_j^{(l)} > a_{nl} y r_l\} \to 0$$

as $n \to \infty$. Using the generalized Lebesgue dominated convergence theorem concludes the proof. $\square$

Conjecture 6.2. *Under the hypothesis of Conjecture 6.1, suppose that $k, l \in \mathcal{D}_3$, $i \in \{1, \ldots, d_k\}$ and $j \in \{1, \ldots, d_l\}$. Let $\{a_{nk}\}$ and $\{a_{nl}\}$ be sequences of positive constants such that for all $x > 0$, $n \Pr(1/(R_k S_i^{(k)}) > a_{nk} x) \to x^{-1}$ and $n \Pr(1/(R_l S_i^{(l)}) > a_{nl} x) \to x^{-1}$ as $n \to \infty$. Then for all $x, y \in \mathbb{R}_+$,*

$$\lim_{n \to \infty} n \Pr\{1/(R_k S_i^{(k)}) > a_{nk} b_k^{1/\rho_k} x, 1/(R_l S_j^{(l)}) > a_{nl} y\} = 0 \; .$$

Chapter 7

Data applications

This chapter contains two applications of the models and methods developed in this thesis to precipitation datasets. This data was kindly provided by Météo France, for which I am very grateful. In Section 7.1, the semiparametric estimation procedure for Archimax copulas, as introduced in Chapter 5, is applied to monthly maxima of daily precipitation for three stations in French Brittany. The strengths of the Archimax model are shown through this illustrative application, and are further pointed out via a small comparative simulation study. Section 7.2 studies a much larger precipitation dataset, weekly maxima for 155 stations spread over metropolitan France. Here, the heterogeneity of the data discourages the use of a single Archimax copula model, so we instead turn to the clustered Archimax copula presented in Chapter 6 which will also allow to model asymptotic independence between stations that are far apart. After discussing certain model choices and implications, a method for finding appropriate clusters is proposed, using an established algorithm equipped with a distance which is tailored to the model. In the second part of Section 7.2, possible directions for modeling joint risk of precipitation at the medium level are discussed.

7.1 Precipitation over French Brittany

In this section, the practical usefulness of the proposed estimation procedure for simple Archimax copula models is illustrated in the context of precipitation monitoring. The data is a trivariate sample of daily precipitation amounts in French Brittany from 1976 to 2016 provided by Météo France. To avoid seasonality, the series is restricted to September to February, during which most extreme events occur. The position of the three stations Belle-Ile, Groix, and Lorient is shown in the left panel of Figure 7.1.

To remove time dependence, and since our primary focus is on extreme precipitation, we considered monthly maxima at each station, totalling 240 observations. Blocking the data by months also eliminates ties; in particular, it avoids the large number of zeros in the sample of daily maxima. This series shows no departures from stationarity; the Ljung and Box–Pierce tests do not reject the hypothesis of temporal independence except at

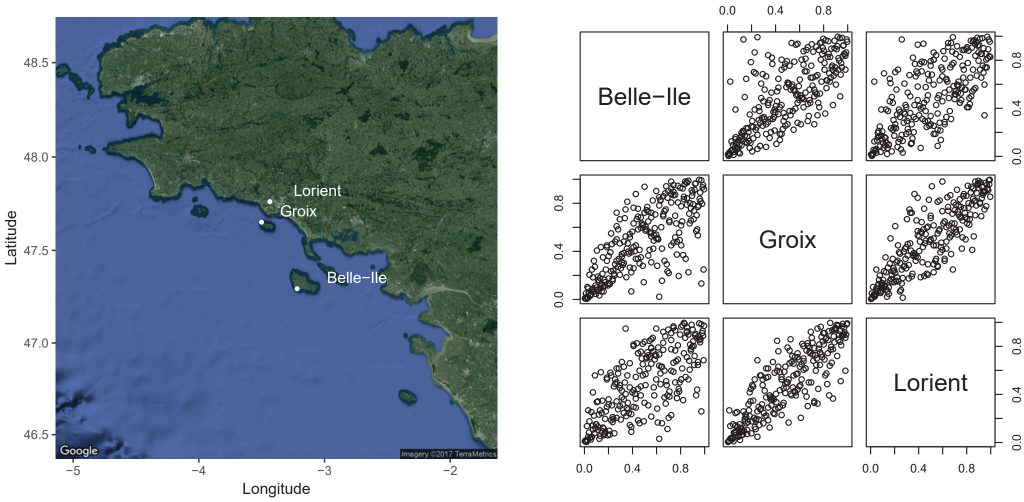

Figure 7.1: Satellite map of French Brittany, showing the sites Belle-Ile, Groix, and Lorient (left). Rankplots of monthly maximum precipitation for the months of September to February, from 1976 to 2016 (right).

Groix, where there is slight evidence of dependence at lags 1 and 2. As the asymptotic results hold for alpha-mixing sequences, time dependence is allowed.

The pairs of the normalized component-wise ranks of monthly maxima are displayed in the right panel of Figure 7.1. These plots show strong correlation between Lorient and Groix, which is not surprising given their geographical proximity. Also apparent is asymmetry between Belle-Ile on the one hand and both Lorient and Groix on the other, in the sense that large precipitation amounts at Groix correspond to large precipitation amounts at Belle-Ile, but not necessarily vice versa, and similarly for Lorient.

Because the data at hand are monthly maxima, one might first think of fitting an extreme-value copula model. However the test of Kojadinovic et al. (2011) clearly rejects the hypothesis that the underlying copula is an extreme-value copula ($p \approx 5 \times 10^{-5}$). This may be explained by the presence of lower-tail dependence, which manifests itself by the clumping of points in the bottom-left corner of the rankplots in the right panel of Figure 7.1. The empirical estimates of the tail probabilities plotted against q in the bottom row of Figure 7.2 also indicate that λ_L in (2.3) for all pairs is likely greater than 0. This phenomenon is not present in multivariate extreme-value distributions, whose pair-wise lower tail dependence coefficients are 0. Archimax copula models advocated in this paper may capture lower-tail as well as extremal dependence. The Clayton-Archimax model is particularly well suited. The latter assumes continuous marginals and an Archimax copula of the form $C_{\psi_\theta,A}$, where A is an arbitrary Pickands dependence function and ψ_θ is the Clayton generator given in Table 4.1. Because ψ_θ for any $\theta > 0$ satisfies Condition 3.2 (a) with $s = 1/\theta$, λ_L of each bivariate margin of $C_{\psi_\theta,A}$ equals $\{2A(1/2)\}^{-1/\theta}$. Furthermore, Condition 3.1 holds with $m = 1$, so that $C_{\psi_\theta,A}$ is in the domain of attraction of the extreme-

Table 7.1: Pair-wise estimates of θ along with 90% asymptotic confidence intervals in the Clayton-Archimax model, model-based estimates of pair-wise Kendall's tau of $C_{\psi_{\theta_n},\hat{A}}$ in the Clayton-Archimax model, and empirical estimates τ_n of pair-wise Kendall's tau.

	$\theta_{n,jk}$	90% C.I.	$\tau(C_{\psi_{\theta_n},\hat{A}})$	τ_n
Belle-Ile & Groix	1.58	(0.77, 2.39)	0.54	0.56
Belle-Ile & Lorient	1.08	(0.49, 1.67)	0.51	0.52
Groix & Lorient	1.27	(0.54, 1.99)	0.64	0.67

value copula C_A. The Clayton-Archimax model is fitted to the data in Section 7.1.1; comparisons with other estimators of the limiting A are considered in Section 7.1.2.

7.1.1 Fitting the Clayton-Archimax model

We begin by estimating the Clayton distortion using the moment-based method presented in Section 5.1. The pair-wise estimates of θ are given in Table 7.1, along with 90% confidence intervals. Because these intervals overlap, there is no evidence against a trivariate Clayton-Archimax model with a common value of θ. The latter is estimated by the average of the pair-wise estimates to be $\theta_n = 1.31$.

The next step consists of estimating A. We use the CFG-type estimator $\hat{A}_{n,c}^{\text{CFG}}$ given in (5.11) with ψ replaced by ψ_{θ_n}. The Pickands-type estimator is not well suited here, because for the estimated value of θ, $s \approx 0.76 < 2$, so that the requirements of Theorem 5.2 are likely not met. In contrast, assuming that Condition 3.3 holds, the assumptions of Theorem 5.1 are fulfilled; Conditions 5.1–5.7 are validated in Example 5.3. Comparing the limiting processes in Theorems 4.1 and 5.1, the additional uncertainty stemming from estimating θ clearly has an impact on the variability of the estimator. To assess the latter in finite samples, we run a pilot simulation which is detailed in Section 7.1.2 and the results of which are shown in Figure 7.4. The boxplots AXC(1) and AXC(2) summarize the IRAE when ψ is known and estimated parametrically, respectively. Unsurprisingly, parameter uncertainty increases the variability of the estimator.

A contour plot of $\hat{A}_{n,c}^{\text{CFG}}$ is shown in the left panel of Figure 7.3. The contour levels of $\hat{A}_{n,c}^{\text{CFG}}$ show a clear global asymmetry, but axial symmetry with respect to Belle-Ile. This pattern corroborates what was seen on the rankplots in Figure 7.1. This asymmetry may be explained by the fact that Belle-Ile is located far off shore. This can lead to strong localized rainfall which does not affect the stations at Groix and Lorient. Although Groix is also an island, it lies much closer to the coast, and is hence not affected by the localized rainfall phenomenon. Furthermore, it can also be seen from pressure maps and radar images that heavy rainfall at Groix and Lorient is mainly due to large-scale weather systems that affect Belle-Ile as well.

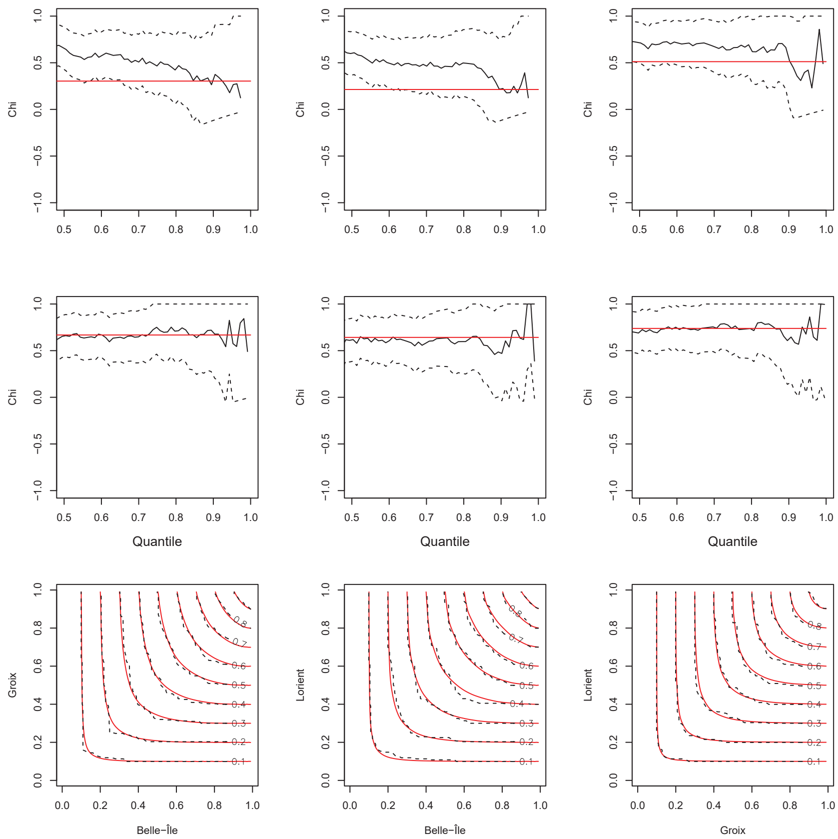

Figure 7.2: Empirical estimates of $\chi_U(q)$ (top) and $\chi_L(q)$ (middle) plotted against q (Quantile) along with 95% confidence bands (black). The red lines indicate the model-based estimates of λ_U (top) and λ_L (middle). Contour plots (bottom) of the empirical copula (black dashed) and the fitted Clayton-Archimax copula (red). The plots correspond to Belle-Ile & Groix (left), Belle-Ile & Lorient (middle), and Groix & Lorient (right).

Finally, we check the fit of the Clayton-Archimax model. Because $\hat{A}_{n,c}^{\text{CFG}}$ is nonparametric, no existing formal goodness-of-fit test for copula models can be used. However, the contours of the fitted trivariate Clayton-Archimax copula seem fairly close to the empirical copula, as evidenced by the bottom panel of Figure 7.2. We also compared various sample dependence measures to their model estimates. To assess the fit in the tails, we consider each pair of stations $j \neq k$, say. Following Coles et al. (1999), we plot the empirical estimates of

$$\chi_U(q) = 2 - \log[\Pr\{F_j(X_j) < q, F_k(X_k) < q\}]/\log(q)$$
$$\chi_L(q) = 2 - \log[\Pr\{F_j(X_j) > 1-q, F_k(X_k) > 1-q\}]/\log(q),$$

against q together with the model-based estimates of the lower and upper tail dependence coefficients λ_L and λ_U for that pair, respectively. To compute the latter, we use that in a bivariate Clayton-Archimax model, as

$$\lambda_L = \lim_{q\to 1} \chi_L(q) = \{2A(1/2)\}^{-1/\theta}, \quad \lambda_U = \lim_{q\to 1} \chi_U(q) = 2 - 2A(1/2).$$

The top two panels of Figure 7.2 show that the model-based estimates approximate the empirical probabilities quite nicely when $q \to 1$, which indicates a good fit in the tails. The contour plots of the empirical copula and the fitted Clayton-Archimax model displayed in the bottom panel of the same Figure match nicely as well. Finally, we compared empirical estimates of pair-wise Kendall's tau with model-based estimates. To compute the latter, we used (4.20) with $\tau_\psi = \theta/(\theta+2)$ and $\tau(A) = \int_0^1 [\{t(1-t)\}/A(t)] dA'(t)$, and approximated the integral in the expression for $\tau(A)$ with finite differences. Table 7.1 shows that the empirical and model-based estimates are very close. Overall, the fit of the Clayton-Archimax model seems adequate, and allows to model the dependence in this trivariate precipitation dataset, not only in extremes, but also in a medium size regime.

7.1.2 Comparison with other estimators of A

If the objective is to specifically assess the joint risk of extreme precipitation, then the estimation of the Pickands dependence function A of the extreme-value attractor of the distribution of the monthly maxima at the three stations is of interest. Because the Clayton-Archimax copula $C_{\psi,A}$ is in the domain of attraction of C_A, the estimator $\hat{A}_{n,c}^{\mathrm{CFG}}$ calculated in the preceding section is also an estimate of the limiting Pickands dependence function. As such, it can be compared to other nonparametric estimators considered in the literature.

The first idea would be to block the data by seasons and consider the maxima over the period from September to February. This reduces the sample size to $n = 40$, but the hypothesis that the underlying copula is an extreme-value copula is no longer rejected by the test of Kojadinovic et al. (2011) ($p \approx 0.43$). Consequently, the multivariate rank-based CFG estimator of Gudendorf and Segers (2012) can be used. Another option would be to use nonparametric estimators of A that only assume that the underlying copula is in the domain of attraction of C_A. We consider the FHM and EKS estimators of Fougères et al. (2015) and Einmahl et al. (2017), respectively. The FHM estimator is denoted as $\mathring{L}_{\mathrm{agg}}$ in Section 5.1 of Fougères et al. (2015), built from Eq. (15) therein, and its tuning parameters are $\kappa_n = 239, a = 0.8, r = 0.8, k_\rho = 237$. The bias-corrected EKS estimator is denoted $\bar{\ell}_{n,k,k_1}$ and its parameters were set to the default choices from the R package `tailDepFun`.

The three competing estimators CFG, FHM, and EKS are displayed in Figure 7.3 along with $A_{n,c}^{\mathrm{CFG}}$ from Section 7.1.1. The contours of the CFG estimator are rougher, which is not surprising given that it is based on 40 observations. Although we expect this estimator to be more variable because it is based on a smaller sample, it is comforting that it shows a similar pattern as $\hat{A}_{n,c}^{\mathrm{CFG}}$; this further confirms that the Clayton-Archimax model is adequate for the data at hand. The contours of the FHM and EKS estimators are much more irregular which makes the plots difficult to interpret.

To compare these estimators further, we ran a pilot simulation study mimicking the data. We generated $N = 1000$ samples of size $n = 240$ from a trivariate Clayton-Archimax copula with $\theta = 1.31$ and the scaled negative extremal Dirichlet Pickands dependence function parameters $\boldsymbol{\alpha} = (1, 2, 3)$ and $\rho = 0.9$ whose shape roughly resembles $\hat{A}_{n,c}^{\mathrm{CFG}}$; see the left panel of Figure 7.4. For each sample, we estimated A by: (i) the CFG-type estimator from (4.4) assuming ψ known; (ii) the CFG-type estimator from (4.4) with θ estimated by the moment estimator θ_n from Section 7.1.1; (iii) the CFG estimator of Gudendorf and Segers (2011) based on block maxima with 40 blocks; (iv) the FHM estimator of Fougères et al. (2015); (v) the EKS estimator of Einmahl et al. (2017). The boxplots of the IRAE are shown in Figure 7.4. Even if ψ is estimated by ψ_{θ_n}, $\hat{A}_{n,c}^{\mathrm{CFG}}$ is superior to the CFG, FHM and EKS estimators especially in terms of bias.

To sum up, this application on precipitation data demonstrates the feasibility of the proposed inference techniques but more importantly illustrates the potential of Archimax copulas to model joint risk in subasymptotic settings. Since the max domain of attraction of Archimax copulas is known, one can check the performance of the latter model by comparing it to models using the max-stable assumption. In this particular data application, the Archimax model accurately captures the bulk and both tails of medium to high precipitation observations. Performance at extreme levels is no doubt also due to the fact that the studied weather stations are located in a relatively small area. To model extremes over larger spatial scales however, more flexible models than those studied herein are required in order to capture asymptotic independence, as noted, e.g., by Huser et al. (2017) and Wadsworth et al. (2017).

7.2 Precipitation over France

This section is concerned with a much larger data set than the one studied in Section 7.1. Here, we have access to precipitation measurements from Météo France at 155 stations across France, for the years 1976 to 2016. As seen in Figure 7.5, some regions such as Côte d'Azur and Île-de-France (the Parisian metropolitan area) feature higher concentrations

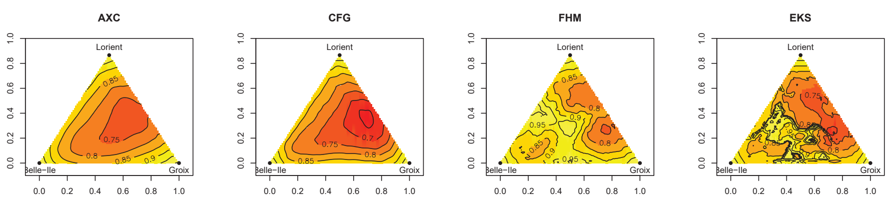

Figure 7.3: AXC: CFG-type estimator $\hat{A}_{n,c}^{\mathrm{CFG}}$ based on monthly maxima and the Clayton-Archimax model. CFG: Rank-based CFG estimator of Gudendorf and Segers (2011) based on seasonal maxima. FHM and EKS: Estimators of Fougères et al. (2015) and Einmahl et al. (2017) based on monthly maxima.

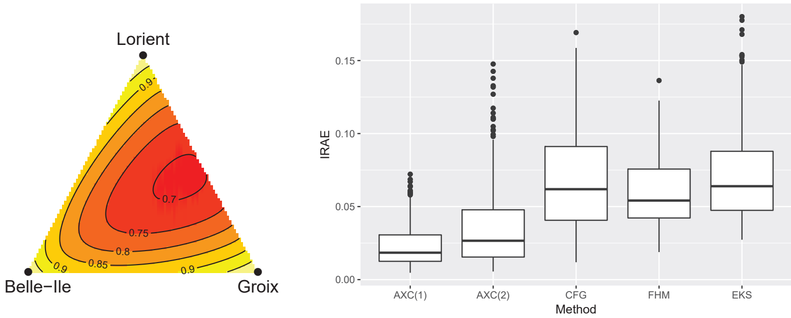

Figure 7.4: Left: **NSD** Pickands dependence function A from Table 4.1 with $\boldsymbol{\alpha} = (1, 2, 3)$ and $\rho = 0.9$. Right: Boxplots of $\mathrm{IRAE}(\hat{A}_n)$ based on $N = 1000$ samples of size $n = 240$ from a 3-variate Clayton-Archimax copula $C_{\psi_\theta, A}$ with $\theta = 1.31$ and the **NSD** A with $\boldsymbol{\alpha} = (1, 2, 3)$ and $\rho = 0.9$. AXC(1): $A_{n,c}^{\mathrm{CFG}}$ from (4.4); AXC(2): $\hat{A}_{n,c}^{\mathrm{CFG}}$ from (5.11) with θ_n from Example 5.2; CFG: the CFG estimator of Gudendorf and Segers (2011) based on block maxima with 40 blocks; FHM: the estimator of Fougères et al. (2015); EKS: the estimator of Einmahl et al. (2017).

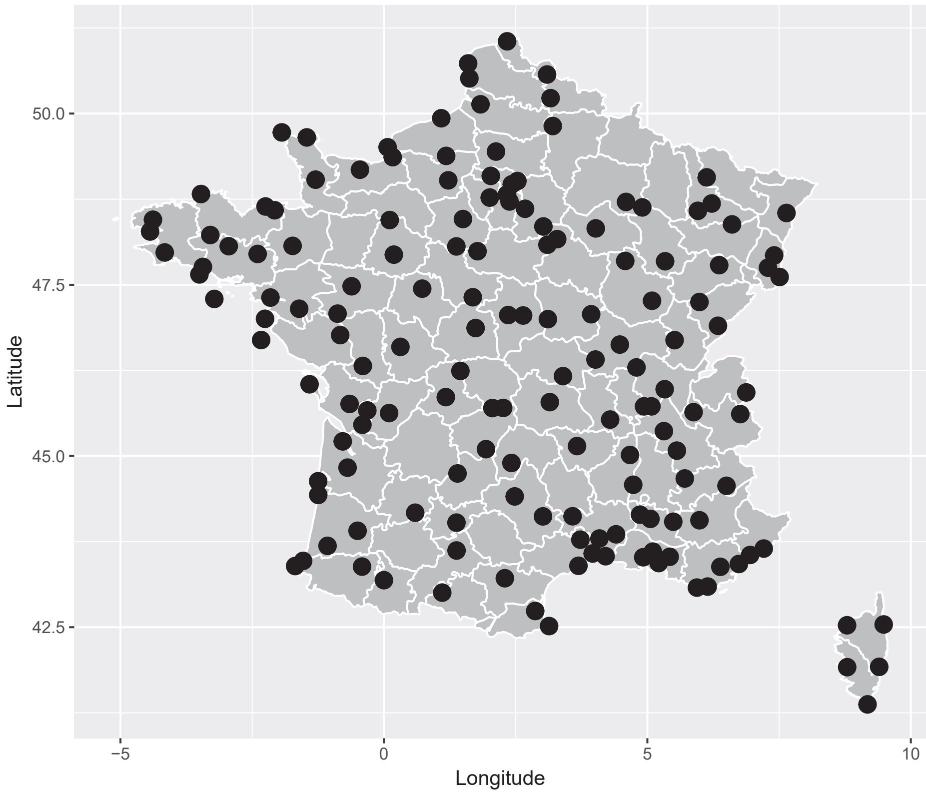

Figure 7.5: Map of the studied 155 weather stations located in metropolitan France.

of stations while other locations are lacking, for example the North-East region of the Ardennes. Since the stations cover a large territory, we need to be more restrictive than in Section 7.1 in order to avoid seasonality in the dataset. As was done in the work of Bernard et al. (2013), which also studied a precipitation dataset with several stations spanning metropolitan France, we restrict the observations to the months of September, October and November. Although the different regions of France can exhibit different weather patterns, this season usually features the heaviest rainfall. For example, the "Orages cévenols", mentioned in the introduction, occur during this time period. Here, the dependence of weekly maxima of daily precipitation measurements is modeled. Since the block size is obviously smaller than the monthly maxima of Section 7.1, an extreme-value copula is not appropriate and lower tail dependence between stations is present as well.

The objective of this section is to fit a clustered Archimax copula $C_{\mathcal{G}\boldsymbol{\psi}\boldsymbol{\ell}}$ proposed and studied in Chapter 6 to this selected dataset. This entails that a partition of the d stations into K sets, denoted $\mathcal{G}$, must be made. Each resulting ordered set $\mathcal{G}_k = \{i_1, \dots, i_{d_k}\}$, $\kappa \in \{1, \dots, K\}$, represents a cluster modeled via an Archimax copula, itself characterized by a stable tail dependence function ℓ_k and an Archimedean generator ψ_k; as such $\boldsymbol{\ell} = \{\ell_1, \dots, \ell_K\}$ and $\boldsymbol{\psi} = \{\psi_1, \dots, \psi_K\}$.

The proposed approach is to first determine an appropriate partition $\mathcal{G}$ via a clustering algorithm presented in Section 7.2.1. Then, an Archimax copula is fit to each cluster using the semiparametric procedure of Chapter 5. As was the case in Section 7.1, the Clayton family was deemed a good choice to model the distortions across all clusters. Of course, in other applications, several distinct Archimedean families for different clusters could be a valid choice. Here, the presence of lower tail dependence made the Clayton family a good candidate; it will be made apparent that a single Archimedean family also greatly simplifies the clustering procedure. One should also note that as shown in Example 6.1, the choice of a Clayton generator for all distortions implies asymptotic independence between all clusters, regardless of the dependence structure of the distortions. Once the Clayton Archimededean generators $\{\psi_{\theta_k}\}_{k=1}^{K}$ are estimated, the stdfs $\{\ell_k\}_{k=1}^{K}$ can also be estimated nonparametrically.

The inter-cluster dependence is modeled through the distortions $(R_1, \dots, R_k)$, which is the topic of Section 7.2.2. The choice of Clayton generators implies that the distortions have marginal densities, and we further assume that the vector $(R_1, \dots, R_k)$ has a parametric copula $C_{\boldsymbol{\xi}}$ with copula density $c_{\boldsymbol{\xi}}$, where $\boldsymbol{\xi} \in \Xi$ for some parameter space Ξ. For example, were we to chose a normal copula for $C_{\boldsymbol{\xi}}$, $\boldsymbol{\xi}$ would be the correlation matrix. Therefore, the density of $(R_1, \dots, R_K)$ is given for all $(r_1, \dots, r_K) \in \mathbb{R}_+^K$, by

$$f_{\boldsymbol{R}}(r_1, \dots, r_K) = c_{\boldsymbol{\xi}}(F_{R_1}(r_1), \dots, F_{R_K}(r_K)) \prod_{k=1}^{K} f_{R_k}(r_k)\ , \tag{7.1}$$

where, as seen in Example 6.1, for each $k \in \{1, \dots, K\}$ and $r_k \in \mathbb{R}_+$,

$$f_{R_k}(r_k) = \frac{\theta_k^d \left\{\prod_{j=0}^{d_k}(1/\theta_k + j)\right\}}{(d_k - 1)!}(1 + \theta_k r_k)^{-1/\theta_k - d_k} r_k^{d_k - 1}$$

and F_{R_k} is the corresponding cumulative distribution function (see Equation (2.6)):

$$F_{R_k}(r_k) = \mathfrak{W}_{d_k}^{-1}\psi_{\theta_k}(r_k) = 1 - \sum_{j=0}^{d_k-2} \frac{(-1)^j r_k^j \psi_{\theta_k}{}^{(j)}(r_k)}{j!} - \frac{(-1)^{(d_k-1)} r_k^{d_k-1} \psi_{\theta_k}{}_{+}^{(d_k-1)}(r_k)}{(d_k - 1)!} .$$

7.2.1 Clustering the stations

In the aforementioned work of Bernard et al. (2013), the partitioning around medoids (PAM) algorithm, introduced by Kaufman and Rousseeuw (1990), is argued to be well suited to cluster asymptotically dependent groups of random variables. To do so, Bernard et al. (2013) use a the PAM algorithm with the F-madogram employed as a distance, viz.

$$d_{ij} = \mathrm{E}\left|F_i(X_i) - F_j(X_j)\right| ,$$

where for each station k, $X_k \sim F_k$ is the random variable of interest. Clearly, the above distance is not affected by marginal behavior and can be seen as being copula-based. As shown by Cooley et al. (2006), the F-madogram is in fact linked to the upper tail dependence coefficient (see Definition 2.3). Indeed, if X_i and X_j have max-stable joint distribution F_{ij} composed of an extreme-value copula C_{ij} with margins F_i and F_j, then

$$d_{ij} = \frac{1}{2}\frac{1 - \lambda_{ij}}{3 - \lambda_{ij}} , \tag{7.2}$$

where $\lambda_{ij} = \lambda(C_{ij})$. In Bernard et al. (2013), it is argued that the PAM algorithm is effective at clustering extremes. Unlike the k-means algorithm which takes averages as cluster centers, the PAM algorithm selects medoids instead, meaning that the distance defined by the equation above remains interpretable at any step of the algorithm.

The dataset of weekly precipitation maxima at hand is clearly not distributed according to a max-stable distribution, as was the case for the monthly maxima of Section 7.1. However, as shown in Proposition 2 p.83 in Murphy (2018), C_{ij} need not be an extreme-value distribution for (7.2) to be a *bona fide* distance. Indeed it suffices that $C_{ij} \in \mathrm{CDA}(C_0)$ for some extreme-value copula C_0.

To compute the distances, we chose to fit the bivariate Clayton-Archimax to all pairs of stations by using the semiparametric approach of Chapter 5. Thus for each $i \neq j$, we have at our disposal an estimate $\hat{\theta}_{ij}$ for the Clayton generator as well as $\hat{\lambda}_{ij} = 2 - \hat{\ell}_{ij}(1, 1)$, recalling that the stdf of the attractor of a Clayton-Archimax copula is equal to that of the Clayton-Archimax copula itself. Note that for these pairwise estimates, zeros were not considered, as is the case in Bernard et al. (2013) (though their threshold is

3mm). Since the objective is to model medium to extreme precipitation and since λ_{ij} in (7.2) is a measure pertaining to the upper tail, removing the zeros removal was deemed acceptable. One can also note that as the model is specified, the dependence between stations that do not belong to the same cluster is not modeled by a bivariate Clayton-Archimax copula. However, the Clayton generator proved to be flexible enough and the estimated tail dependence coefficients were very similar when estimated via two other techniques, a non-parametric estimator (see Figure 7.2) and a parametric approach using a t-copula. An estimator for (7.2) is then simply obtained by plugging in $\hat{\lambda}_{ij}$.

Grouping stations which exhibit strong asymptotic dependence is not sufficient for the clustered Archimax copula model to be applied. For each cluster $k \in \{1, \ldots, K\}$, the assumption of a single distortion R_k affecting the extreme regime of the d_k stations characterized by $(S_1^{(k)}, \ldots, S_{dk}^{(k)})$ also needs to be reasonable. To account for the assumption of a single distortion per cluster, we introduce weights for each pairwise distance, viz

$$d_{ij}^{\mathrm{W}} = w_{ij} \frac{1}{2} \frac{1 - \hat{\lambda}_{ij}}{3 - \hat{\lambda}_{ij}} \,,$$

where

$$w_{ij} = \frac{\sum_{k \notin \{i,j\}} \hat{\lambda}_{ij} \hat{\lambda}_{ik} |\hat{\theta}_{ik} - \hat{\theta}_{jk}|}{\sum_{k \notin \{i,j\}} \hat{\lambda}_{ij} \hat{\lambda}_{ik}} \,.$$

Theses weights encourage stations within the same cluster to have pairwise estimates θ_{ik} that are similar. The product $\hat{\lambda}_{ij}\hat{\lambda}_{ik}$ ensures that for stations k which are "far" (in the sense of extremal dependence) from i and j, the differences $|\hat{\theta}_{ik} - \hat{\theta}_{jk}|$ have less of an impact.

We also mix d^{W} above with the classical euclidean distance betweens stations d^{G}. For two stations i, j with longitudes $\mathrm{lon}_i, \mathrm{lon}_j$ and latitudes $\mathrm{lat}_i, \mathrm{lat}_j$, this distance is simply $d_{ij}^{\mathrm{G}} = \{(\mathrm{lon}_i - \mathrm{lon}_j)^2 + (\mathrm{lat}_i - \mathrm{lat}_j)^2\}^{1/2}$ and the resulting distance $d^\star$ is defined by

$$d_{ij}^\star = (1 - \alpha) \frac{d_{ij}^{\mathrm{W}}}{\max\limits_{i \neq j} d_{ij}^{\mathrm{W}}} + \alpha \frac{d_{ij}^{\mathrm{G}}}{\max\limits_{i \neq j} d_{ij}^{\mathrm{G}}} \,,$$

with $\alpha = 1/3$. This mixing parameter was chosen to be small since the geographic distance d^{G} was employed to avoid rare and spurious groupings of stations that were far apart, and does not represent "climatological" distance well. For example, the east and west coast of Corsica will often observe different weather, even though the distance between them is relatively small. In the case of Corsica, this is explained by the mountainous topography of the island.

For a given number of clusters K, the PAM algorithm, which is implemented in the R package `cluster`, consists in choosing K cluster medoids at random and following three steps:

(1) Each station is assigned to the nearest medoid according to $d_{ij}^\star$.

(2) For each cluster $k \in \{1, \ldots, K\}$, find the new medoid that minimizes the total intra-cluster distances based on $d^\star_{ij}$.

(3) Repeat steps (1) and (2) as long as at least one medoid has changed.

To chose the number of clusters K, we us the average silhouette coefficient as introduced by Rousseeuw (1987) and implemented in the `cluster` package. According to this method, $K = 41$ clusters is a good choice; the result of the PAM algorithm can be seen in Figure 7.6. The mean silhouette coefficient for $k = 41$ is about 0.32. The closer a coefficient is to 1, the better the clustering, while a value of 0 is synonymous for a non-informative clustering.

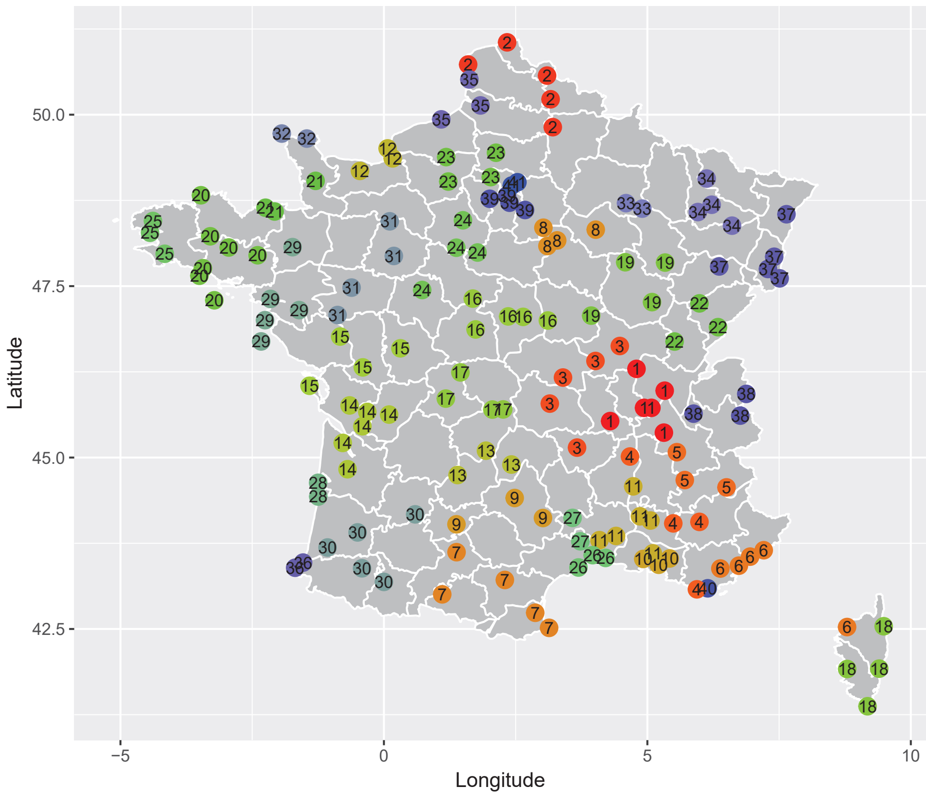

Figure 7.6: Map of the clustered 155 weather stations resulting from the PAM algorithm using $d^\star_{ij}$.

The resulting clustering exhibits interesting cluster shapes. Unfortunately, the 40th cluster is a singleton and cannot be captured by our modeling procedure. One can note that mountainous regions are often more heterogeneous, which can be seen in the south-east and south-west of the territory. Moreover, the 1st and 11th clusters seem to follow the Rhône valley. The Cévennes region, roughly corresponding to cluster 27, is separated from other neighboring stations which is expected.

Once the clusters are determined, each Archimedean generator ψ_k, for $k \in \{1, \ldots, K\}$, can be estimated by averaging the pairwise estimates, viz.

$$\hat{\theta}_k = \frac{2}{d_k(d_k - 2)} \sum_{i,j \in I_k, i \neq j} \hat{\theta}_{ij} \ ,$$

where $I_k = \{j : \text{station } j \text{ is in cluster } k\}$. The stable tail dependence functions $\{\ell_k\}_{k=1}^K$ can then be estimated nonparametrically. For ease of notation, let $\{\hat{\ell}_k\}_{k=1}^K$ and $\{\psi_{\hat{\theta}_k}\}_{k=1}^K$ denote the estimated stdfs and Archimedean generators, respectively.

This clustering procedure is work in progress whose quality needs to be evaluated via simulations. Generating data from the model would allow to test the robustness of the choices made, such as the distance matrix used, the estimation of the λ_{ij} coefficients or the choice of the number of clusters K. The procedure will also be evaluated by applying it to other datasets, such as a portfolio of stocks from various industries. The use of Euclidean distance between stations is an ad-hoc way to avoid the very rare but obvious misclassifications that occurred without it. The drawbacks involve the risk of over-fitting and inducing user bias in the clustering; the objective in the future is to create a robust algorithm that does not require this type of tuning and can be applied in other settings.

7.2.2 Modeling the distortions

At this stage, suppose that $\mathcal{G}$ as well as the functional parameters $\{\psi_k\}_{k=1}^K$ and $\{\ell_k\}_{k=1}^K$ have been estimated. Recall the form of the density of $\boldsymbol{R} = (R_1, \ldots, R_K)$ given in (7.1). The aim of this section is to discuss a strategy on how the parameter $\boldsymbol{\xi}$ of the copula of the distortions $\boldsymbol{R}$ can be estimated. To begin, suppose that $\boldsymbol{U} = (\boldsymbol{U}^{(1)}, \ldots, \boldsymbol{U}^{(K)}) \sim C_{\mathcal{G}\boldsymbol{\psi}\boldsymbol{\ell}}$, so that $(\phi_1\{\boldsymbol{U}^{(1)}\}, \ldots, \phi_K\{\boldsymbol{U}^{(K)}\}) \stackrel{d}{=} (R_1\boldsymbol{S}^{(1)} \ldots, R_K\boldsymbol{S}^{(K)})$. The main difficulty in estimating $\boldsymbol{\xi}$ lies in the fact that even if $\boldsymbol{U}$ were observable, $\boldsymbol{R}$ cannot be observed. However, the following result will prove to be helpful.

Proposition 7.1. *Suppose that $(R_1\boldsymbol{S}^{(1)}, \ldots, R_K\boldsymbol{S}^{(K)})$ is a $d = \sum_{k=1}^K d_k$-dimensional random vector as in Definition 6.1 and assume further that $(R_1, \ldots, R_K)$ has a density $f_{\boldsymbol{R}}$. For all $k \in \{1, \ldots, K\}$, let*

$$B_k = \ell_k(\mathbf{1}) \min_{i \in \{1, \ldots, d_k\}} \{S_i^{(k)}\} \ ,$$

where $\mathbf{1}$ is a vector of 1's of dimension d_k. Then, the density of $\boldsymbol{Y} = (Y_1, \ldots, Y_K) = (R_1B_1, \ldots, R_KB_K)$ is given, for all $(y_1, \ldots, y_K) \in \mathbb{R}_+^K$, by

$$f_{\boldsymbol{Y}}(y_1, \ldots, y_K) = \mathrm{E}\left(\frac{f_{\boldsymbol{R}}(y_1/D_1 \ldots, y_K/D_K)}{D_1 \ldots D_K}\right) \ , \tag{7.3}$$

where $D_1, \ldots, D_K$ are independent and such that $D_k \sim B(1, d_k - 1)$ for $k \in \{1, \ldots, K\}$.

Proof. The distribution of B_k can be seen to be Beta $B(1, d_k - 1)$:

$$\begin{aligned}\Pr(B_k > s) &= \Pr(S_1^{(k)} \wedge \ldots \wedge S_{d_k}^{(k)} > s/\ell_k(\mathbf{1})) \\ &= \Pr(S_1^{(k)} > s/\ell_k(\mathbf{1}), \ldots, S_{d_k}^{(k)} > s/\ell_k(\mathbf{1})) = (1-s)^{d_k - 1} .\end{aligned}$$

To obtain the density of $\boldsymbol{Y}$, we simply apply the transformation theorem. To do so, define the said transformation t as follows:

$$t : (R_1, \ldots, R_K, B_1, \ldots, B_K) \longmapsto (R_1 B_1, \ldots, R_K B_K, B_1, \ldots, B_K) .$$

Thus, the components of the inverse of t are as follows:

$$t_j^{-1}(y_1, \ldots, y_K, y_{K+1}, \ldots, y_{2K}) = \begin{cases} y_j / y_{j+K} & \text{for} \quad j \in \{1, \ldots, K\} \\ y_j & \text{for} \quad j \in \{K+1, \ldots, 2K\} \end{cases}$$

The Jacobian is then an upper triangular matrix with determinant equal to $\prod_{j=1}^{K} (1/y_{j+K})$. Therefore, the density of $(R_1 B_1, \ldots, R_K B_K, B_1, \ldots, B_K)$ is given, for $(y_1, \ldots, y_K) \in \mathbb{R}_+^K$ and $(y_{1+K}, \ldots, y_{2K}) \in (0,1)^K$, by

$$f_{\boldsymbol{R}}(y_1/y_{1+K}, \ldots, y_K/y_{2K}) \prod_{j=1}^{K} \frac{(1 - y_{j+K})^{d_k - 2}}{(d_k - 1) y_{j+K}} .$$

To get to the density of $(R_1 B_1, \ldots, R_K B_K)$, we must integrate out the second half of the vector:

$$\int_0^1 \cdots \int_0^1 f_{\boldsymbol{R}}(y_1/y_{1+K}, \ldots, y_K/y_{2K}) \prod_{j=1}^{K} \frac{(1 - y_{j+K})^{d_k - 2}}{(d_k - 1) y_{j+K}} dy_{j+K} .$$

This is equal to

$$\mathrm{E}\left(\frac{f_{\boldsymbol{R}}(y_1/D_1 \ldots, y_K/D_K)}{D_1 \ldots D_K} \right) .$$

□

If the density of is the form (7.1), then (7.3) only depends on the parameter $\boldsymbol{\xi}$. One can therefore attempt estimating $\boldsymbol{\xi}$ using maximum likelihood techniques. However, maximizing the likelihood based on (7.3) directly is unwieldy since the expectation cannot be calculated explicitly. Instead, we propose to use an EM algorithm. To this end, suppose for the sake of simplicity that $\boldsymbol{U}_1, \ldots, \boldsymbol{U}_n$ is a sample from $C_{\mathcal{G}\boldsymbol{\psi}\boldsymbol{\ell}}$ where $\mathcal{G}$, $\boldsymbol{\psi} = \{\psi_k\}_{k=1}^K$ and $\boldsymbol{\ell} = \{\ell_k\}_{k=1}^K$ are all known. Suppose also that the $\boldsymbol{\psi} = \{\psi_k\}_{k=1}^K$ are Clayton, hence their parameters $\{\theta_k\}_{k=1}^K$ are also known. Then, let $\boldsymbol{Y}_1, \ldots, \boldsymbol{Y}_n$ be defined, for all $i = 1, \ldots, n$, as follows:

$$\boldsymbol{Y}_i = (Y_{i1}, \ldots, Y_{iK}) = \left(\frac{\min\limits_{j \in \{1, \ldots, d_1\}} \{\phi_{\theta_1}(U_{ij}^{(1)})\}}{\ell_1(\mathbf{1})}, \ldots, \frac{\min\limits_{j \in \{1, \ldots, d_K\}} \{\phi_{\theta_K}(U_{ij}^{(K)})\}}{\ell_K(\mathbf{1})} \right) , \tag{7.4}$$

making it a random sample from (7.3). In the above display, $U_{ij}^{(k)}$ is the i-th copula observation of the j-th component of the cluster k. To devise the EM algorithm, write the log likelihood as follows:

$$\ln L(\boldsymbol{\xi}; \boldsymbol{Y}_1, \ldots, \boldsymbol{Y}_n, ; \boldsymbol{B}_1, \ldots, \boldsymbol{B}_n) = \sum_{i=1}^{n} \ln f_{\boldsymbol{X}|\boldsymbol{B}}(\boldsymbol{X}_i | \boldsymbol{B}_i; \boldsymbol{\xi}) + \sum_{i=1}^{n} \ln f_{\boldsymbol{B}}(\boldsymbol{B}_i) \, .$$

Conveniently, the second part in the above does not depend on $\boldsymbol{\xi}$ and can be dropped. Therefore,

$$\ln L(\boldsymbol{\xi}; \boldsymbol{Y}_1, \ldots, \boldsymbol{Y}_n, ; \boldsymbol{B}_1, \ldots, \boldsymbol{B}_n) \propto \sum_{i=1}^{n} \ln f_{\boldsymbol{X}|\boldsymbol{B}}(\boldsymbol{X}_i | \boldsymbol{B}_i; \boldsymbol{\xi}) \, .$$

Recalling the form of the density in (7.1), one can further simplify the above to obtain the form

$$\begin{aligned} \ln L(\boldsymbol{\xi}; \boldsymbol{Y}_1, \ldots, \boldsymbol{Y}_n, ; \boldsymbol{B}_1, \ldots, \boldsymbol{B}_n) \propto \quad & \ln \tilde{L}(\boldsymbol{\xi}; \boldsymbol{Y}_1, \ldots, \boldsymbol{Y}_n, ; \boldsymbol{B}_1, \ldots, \boldsymbol{B}_n) \\ = & \sum_{i=1}^{n} \ln c_{\boldsymbol{\xi}}(F_{\boldsymbol{R}}(\boldsymbol{Y}_i / \boldsymbol{B}_i)) \, , \end{aligned}$$

where the marginal densities of $\boldsymbol{R}$ were removed as they do not depend on $\boldsymbol{\xi}$. The E step at time step s consists in computing the conditional expectation to define the following objective function:

$$Q(\boldsymbol{\xi}; \boldsymbol{\xi}^{(k)}) = \mathrm{E}\left[\ln \tilde{L}(\boldsymbol{\xi}; \boldsymbol{Y}_1, \ldots, \boldsymbol{Y}_n, ; \boldsymbol{B}_1, \ldots, \boldsymbol{B}_n) | \boldsymbol{Y}_1, \ldots, \boldsymbol{Y}_n; \boldsymbol{\xi}^{\{s\}} \right] \, ,$$

where $\boldsymbol{\xi}^{\{s\}}$ denotes the parameter estimate at time step s. The M step is then to maximize Q with respect to $\boldsymbol{\xi}$, viz.

$$\boldsymbol{\xi}^{\{s+1\}} = \arg\max_{\boldsymbol{\xi}} Q(\boldsymbol{\xi}; \boldsymbol{\xi}^{(s)}) \, .$$

The E and M steps are repeated until convergence ensues. To perform the E step, the expectation needs to be approximated via Monte Carlo. This is done by drawing from the distribution of $\boldsymbol{B}|\boldsymbol{Y}$, whose density is proportional to $f_{\boldsymbol{Y},\boldsymbol{B}}$. These draws can be performed using either importance sampling or rejection sampling, which is currently being investigated.

Of course, the properties of the resulting estimator $\hat{\boldsymbol{\xi}}$ need to be investigated, both theoretically and via simulations; this is the objective of the immediate future. In fact, the matter is further complicated by the fact that we have to resort to using pseudo-observations. The first level of approximation is that the copula sample is in fact a rank-based pseudo sample as given in Equation (2.13). Moreover, the inverse generators $\{\phi_k\}_{k=1}^{K}$ are estimated parametrically and the stdfs $\{\ell_k\}_{k=1}^{K}$ are estimated nonparametrically. Thus (7.4) is replaced by the following, for $i = 1, \ldots, n$:

$$\left(\frac{\min_{j \in \{1,\ldots,d_1\}} \{\phi_{\hat{\theta}_1}(\hat{U}_{ij}^{(1)})\}}{\hat{\ell}_1(\mathbf{1})}, \ldots, \frac{\min_{j \in \{1,\ldots,d_k\}} \{\phi_{\hat{\theta}_k}(\hat{U}_{ij}^{(k)})\}}{\hat{\ell}_k(\mathbf{1})}, \ldots, \frac{\min_{j \in \{1,\ldots,d_K\}} \{\phi_{\hat{\theta}_K}(\hat{U}_{ij}^{(K)})\}}{\hat{\ell}_K(\mathbf{1})} \right) ,$$

where $\hat{U}_{ij}^{(k)}$ is the normalized rank of the i-th observation of the j-th component of the cluster k. The pseudo-observations in the above display therefore inherit uncertainty due to the estimation of the copula sample, the partition $\mathcal{G}$, the generators $\{\psi_k\}_{k=1}^K$ and the stdfs $\{\ell_k\}_{k=1}^K$. In the near future, this should be further investigated through simulation studies.

Chapter 8

Conclusion and future work

The first objective of my Ph.D. was to develop inference techniques for the Archimax class of copulas. My understanding is that this family, introduced in the bivariate setting by Capéraà et al. (2000), was mostly seen as a tool for simulation studies. Indeed, one can test the effectiveness of estimation techniques for asymptotic dependence structures using a variety of Archimax copulas with known extreme-value attractor. In Fougères et al. (2015), the proposed inference procedure for limiting stable tail dependence functions only assumes the existence of an extreme-value attractor and involves the choice of a threshold. The simulation study therein uses Archimax copulas to study the finite sample performance since a variety of asymptotic regimes can be tested. In Bücher et al. (2019), the efficiency of the block maxima and peaks over threshold methods are compared in the multivariate setting. Through second order methods, the authors find that the convergence of one method usually implies the convergence of the other; however the rates might be different depending on the underlying copula. The Archimax family is employed in this paper to illustrate this fact both theoretically and through a simulation study. Depending on the choice of the Archimedean generator and its index of regular variation at zero, either the block maxima or the peaks over threshold method will prove to be asymptotically superior. This result is particularly interesting given the preference, in recent years, for the latter method in the extreme value analysis community.

While there is still work to be done to improve its ease of use, I believe that the Archimax family also has its place in applications to risk modeling, in areas ranging from insurance to environmental sciences. As seen in Chapter 7, it appears that this family can be well suited to fit multivariate datasets which are not "yet" distributed according to an extreme-value distribution. Taking large block sizes or imposing high thresholds can be quite costly; using an Archimax model allows for the retention of a greater proportion of the data. The estimation procedure proposed in this thesis is geared toward inference on the extremal dependence regime of the data at hand. Since the main driver of the said regime is the stable tail dependence function, a nonparametric approach offers a certain flexibility and granularity. This of course comes with typical drawbacks, for example the fact that the estimator itself is not a valid stdf and the fact that goodness-of-fit is hard to check. The parametric estimation of the Archimedean generator comes with several advantages. Firstly, many single parameter families have already been extensively studied in the literature. With some exploratory data analysis, one can identify certain properties that seem to be present and choose an appropriate Archimedean family, as is done in Chapter 7. Since the indices of regular variation at zero and infinity of the generators are often linked to their parameter, this allows to estimate the maximum and minimum attractor of the (assumed) underlying Archimax copula, when combined with the estimated stdf. Finally, the assumption of a parametric family for the Archimedean generator made the extension of the weak convergence results presented in Chapter 5, more manageable. Three other estimation procedures were considered during the course of my Ph.D. but did not make it into this thesis, nor the resulting paper. The first was a completely parametric approach; but having to choose a family for the stdf was not ideal, especially in higher dimensions where asymmetry often implies many extra parameters to estimate. However, likelihood-based estimation boasts many advantages and this is an option worth having in my opinion. Secondly, I attempted a pairwise semi-parametric approach where the stdf and Archimedean generator were iteratively estimated assuming the knowledge of the other functional parameter until some stability was attained. The estimator of the stdf was the same as the one proposed in this thesis while the estimator of the Archimedean generator was based on inverting Kendall's tau. In simulations, this procedure would sometimes diverge and theoretical grounding to study the method was lacking. Finally, I attempted a completely non-parametric estimation of the Archimedean generator using the nested diagonal property also present in Archimedean copulas. This is in fact an extension of the work of Di Bernardino and Rulliere (2013), but results of the procedure in simulations were not encouraging. Thus the semiparametric approach was retained and extensively studied via asymptotics, simulations and an illustrative application to a trivariate rainfall dataset. This work makes up for most of this thesis and resulted in the paper titled "Inference for Archimax copulas" to be published in the *Annals of Statistics* this year. I also plan to write an R package in the upcoming months to make the tools developed for Archimax models available online.

Given the promising results for the simple Archimax model, it appeared natural to extend it to a hierarchical construction as done in Chapter 6. For two univariate margins of a distribution with an Archimax copula to be asymptotically independent, their marginal bivariate copula must necessarily be Archimedean. To avoid this restriction and to allow for a more parsimonious model in higher dimensions, Archimax copulas can be linked together via a dependence structure on their distortions, giving rise to the clustered Archimax copula. In Chapter 6, the maximum domain of attraction of a clustered Archimax copula is found and shown to have certain desirable properties. Namely, extreme-dependence between clusters is found to be very flexible, due to both the distortions and their own asymptotic dependence having an impact. Notably, an interesting by-product of this work is the discovery of a new way to construct stable tail dependence functions. I hope to complete the proof of the extension proposed in Section 6.4 which would cover virtually all possible cases of distortions and their attractors. Additionally, determining the minimum domain of attraction of a clustered Archimax copula appears to be an achievable goal in the near future, most likely by employing similar techniques to those used in Section 6.3. These results put together will make up for a paper on clustered Archimax copulas that I am expecting to submit in the upcoming months.

Section 7.2 contains work in progress for a paper in preparation with Samuel Perreault from Université Laval. The objective is to develop tools to use clustered Archimax copulas in an applied setting. To do so, we are working with two datasets. The first, as presented in the aforementioned section, is a dataset consisting of precipitation measured at over one hundred stations spread across a large territory. Here we want to identify small regions which have high risk of joint extreme precipitation, with their shape describing storm patterns during the studied season. The second, which we are currently working with and therefore did not make into this thesis, is a portfolio dataset consisting of stock returns where one can easily imagine different groups of stocks in the same industry being asymptotically dependent, while stocks from different industries might be less intertwined. Our approach so far has been to adapt existing clustering algorithms to the model in order to find groups of variables with strong asymptotic dependence, and for whom the assumption of a single distortion affecting the said extreme regime is a reasonable assumption to make. The second step is then to fit Archimax copulas to each cluster with the semiparametric approach discussed earlier. Finally, the dependence between the distortions is inferred upon. This sequential approach clearly has the drawback of not taking into account the uncertainty of the clustering when fitting the Archimax copulas and the distribution of the distortions. I think an interesting project would be to borrow from bayesian methodology to improve on this, as is done, for example, in Vettori et al. (2019).

From the work I have done during my Ph.D., three problems for future research have become apparent to me. Although they are related to this thesis, they are more ambitious than the extensions and improvements already suggested in the previous paragraphs. The first one is related to the stochastic representation of simple and clustered Archimax copulas. The representation used in this thesis is the most general as it works for any Archimedean generator, while the frailty representation (see Section 4.2 in Joe (2014), or McNeil (2008)) is only valid for completely monotone generators. However, this generality comes at the expense of handiness. It is not known how to simulate from the random vector $\boldsymbol{S}$ of $R\boldsymbol{S}$ in (2.11) apart for some examples such as the logistic stdf. However, there appears to be some parallels to be drawn from the spectral representation of max-stable processes of De Haan et al. (1984). It seems like the Poisson process in the latter can be replaced by a binomial process whose parametrization depends on the stable tail dependence function that characterizes $\boldsymbol{S}$, however many complications arise. This would amount to sampling from any multivariate extreme dependence structure and to the best of my knowledge, this is a high-reaching problem that has been sought after for some years. The second long term project I have in mind stems from the precipitation dataset I have available from Météo France. While the multivariate approach from this thesis can answer some questions, there is an obvious flaw in that inference at locations between stations is not available. The spatial nature of the data calls for a spatial model, and one could argue that the temporal aspect should also be modeled. In the current literature, a lot of effort is being put into developing models that are able to capture the extremal dependence structure of precipitation (see, for example, Huser et al. (2017), Wadsworth et al. (2017) and Bacro et al. (2019)). Indeed, as previously discussed, there is evidence that rainfall quickly loses its extreme dependence as distance between locations grows. Can the distortion on an extreme regime paradigm be applied in a spatial setting? While there are similarities with random scale mixtures, the work done in Chapter 6 points toward the distortions being themselves a field over the observed domain. The margins would no longer be Archimax copulas, as it is the case with inter-cluster margins in the clustered Archimax copula model, but perhaps this could be an interesting area to explore. Finally, causality and extremes has garnered a lot of interest recently, especially in applications related to climate change research. In the data applications presented in this thesis, storm patterns were picked up both by the asymmetry in the stable tail dependence and the shapes of the regions identified as asymptotically dependent clusters of stations. This constitutes, in my opinion, a very compelling reason to explore how concepts in extremal dependence can be linked to causality.

Appendix A

Detailed simulation study results

This section contains the detailed results of the simulation study from Section 4.3 in the form of tables containing the means of errors obtained from 1000 Monte Carlo replicates.

Tables A.12, A.13 and A.14 compare results for logistic (**LG**) and discrete spectral measure-type (**DSM**) Pickands dependence functions. Following the notation of Fougères et al. (2013), the parameter choices for the latter are provided below. We have $m = 10$ and $\boldsymbol{w}^{(d)}$ the matrix of weight parameters, where d denotes the dimension.

$$\boldsymbol{w}^{(2)} = \begin{bmatrix} 1.00 & 0.93 & 0.87 & 0.80 & 0.73 & 0.67 & 0.60 & 0.53 & 0.47 & 0.40 \\ 0.00 & 0.07 & 0.13 & 0.20 & 0.27 & 0.33 & 0.40 & 0.47 & 0.53 & 0.60 \end{bmatrix}, \tag{A.1}$$

$$\boldsymbol{w}^{(4)} = \begin{bmatrix} 0.67 & 0.00 & 0.33 & 0.33 & 0.00 & 0.33 & 0.33 & 0.00 & 0.00 & 0.00 \\ 0.33 & 1.00 & 0.33 & 0.00 & 0.00 & 0.33 & 0.00 & 0.00 & 0.33 & 0.00 \\ 0.00 & 0.00 & 0.33 & 0.67 & 1.00 & 0.00 & 0.33 & 0.67 & 0.00 & 0.00 \\ 0.00 & 0.00 & 0.00 & 0.00 & 0.00 & 0.33 & 0.33 & 0.33 & 0.67 & 1.00 \end{bmatrix}, \tag{A.2}$$

$$\boldsymbol{w}^{(10)} = \begin{bmatrix} 0.33 & 0.00 & 0.00 & 0.33 & 0.00 & 0.00 & 0.00 & 0.33 & 0.00 & 0.00 \\ 0.33 & 0.00 & 0.00 & 0.00 & 0.00 & 0.33 & 0.00 & 0.00 & 0.00 & 0.00 \\ 0.00 & 0.33 & 0.00 & 0.33 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 \\ 0.00 & 0.33 & 0.67 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 \\ 0.33 & 0.00 & 0.00 & 0.00 & 0.33 & 0.33 & 0.00 & 0.33 & 0.33 & 0.00 \\ 0.00 & 0.33 & 0.00 & 0.00 & 0.00 & 0.00 & 0.33 & 0.00 & 0.00 & 0.00 \\ 0.00 & 0.00 & 0.33 & 0.00 & 0.33 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 \\ 0.00 & 0.00 & 0.00 & 0.33 & 0.33 & 0.00 & 0.33 & 0.00 & 0.33 & 0.00 \\ 0.00 & 0.00 & 0.00 & 0.00 & 0.00 & 0.33 & 0.33 & 0.00 & 0.00 & 0.00 \\ 0.00 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 & 0.33 & 0.33 & 1.00 \end{bmatrix}. \tag{A.3}$$

Table A.1: Average Integrated relative absolute error (IRAEx100) and Integrated squared error (ISEx10000) of $A_{n,c}^{\mathrm{CFG}}$ and $A_{n,c}^{\mathrm{P}}$ for 2-dimensional Archimax copula $C_{\psi,A}$ samples of size $n \in \{200, 500, 1000\}$. The Pickands dependence function A is **LG** with parameter $\varrho = 2$ so that $\tau(A) = 1/2$, where $\tau(A) = \tau(C_A)$ is Kendall's tau of the bivariate extreme-value copula C_A. There are four choices for the Archimedean generator ψ, Clayton, Frank, Gumbel and Joe, each with four parameter choices so that $\tau(\psi) \in \{1/5, 2/5, 3/5, 4/5\}$, where $\tau(\psi) = \tau(C_\psi)$ is Kendall's tau of the bivariate Archimedean copula C_ψ. There are 1000 Monte Carlo replicates.

	n	200				500				1000			
	error	IRAE($\times$100)		ISE($\times$10000)		IRAE($\times$100)		ISE($\times$10000)		IRAE($\times$100)		ISE($\times$10000)	
ψ	$\tau(\psi)$ | estimator	CFG	P	CFG	P	CFG	P	CFG	P	CFG	P	CFG	P
Clayton	1/5	1.39	2.42	1.97	6.08	0.86	1.76	0.75	3.16	0.64	1.33	0.42	1.82
	2/5	1.49	8.3	2.27	90.48	0.92	7.84	0.87	83.79	0.69	7.66	0.48	79.73
	3/5	1.59	28.97	2.68	26.55×10^2	0.98	35.08	1.02	41.92×10^2	0.75	36.38	0.56	44.95×10^2
	4/5	2.25	620.74	5.38	1915×10^6	1.24	86.32×10	1.64	66.08×10^6	0.87	48.34×10	0.77	30.40×10^5
Frank	1/5	1.38	1.84	1.9	3.43	0.84	1.16	0.72	1.33	0.64	0.87	0.41	0.77
	2/5	1.36	2.23	1.89	5.08	0.84	1.5	0.71	2.27	0.62	1.06	0.38	1.14
	3/5	1.31	2.9	1.75	8.73	0.82	1.86	0.68	3.56	0.61	1.39	0.37	1.9
	4/5	1.41	4.35	2.05	20.9	0.81	2.78	0.68	8.01	0.59	2.14	0.35	4.65
Gumbel	1/5	1.43	1.85	2.08	3.48	0.91	1.19	0.84	1.43	0.67	0.88	0.45	0.78
	2/5	1.49	2.27	2.23	5.23	0.95	1.49	0.92	2.25	0.69	1.1	0.48	1.21
	3/5	1.58	3.23	2.56	11.15	0.97	2.25	0.97	5.12	0.72	1.66	0.51	2.74
	4/5	1.97	7.46	4.07	69.08	1.11	5.57	1.27	36.18	0.79	4.32	0.62	20.31
Joe	1/5	1.46	1.79	2.17	3.27	0.92	1.14	0.87	1.33	0.68	0.84	0.46	0.71
	2/5	1.54	2.04	2.4	4.2	0.98	1.32	1	1.74	0.72	0.94	0.51	0.91
	3/5	1.71	2.41	3.1	5.78	1.05	1.56	1.17	2.44	0.77	1.12	0.59	1.31
	4/5	2.37	3.38	5.81	12.03	1.28	2.21	1.7	4.99	0.91	1.56	0.83	2.53

Table A.2: Average Integrated relative absolute error (IRAEx100) and Integrated squared error (ISEx10000) of $A_{n,c}^{\mathrm{CFG}}$ and $A_{n,c}^{\mathrm{P}}$ for 2-dimensional Archimax copula $C_{\psi,A}$ samples of size $n \in \{200, 500, 1000\}$. The Pickands dependence function A is **NSD** with parameters $\boldsymbol{\alpha} = (1, 2)$, $\rho = 0.59$, so that $\tau(A) = 1/2$, where $\tau(A) = \tau(C_A)$ is Kendall's tau of the bivariate extreme-value copula C_A. There are four choices for the Archimedean generator ψ, Clayton, Frank, Gumbel and Joe, each with four parameter choices so that $\tau(\psi) \in \{1/5, 2/5, 3/5, 4/5\}$, where $\tau(\psi) = \tau(C_\psi)$ is Kendall's tau of the bivariate Archimedean copula C_ψ. There are 1000 Monte Carlo replicates.

	n	200				500				1000			
	error	IRAE(×100)		ISE(×10000)		IRAE(×100)		ISE(×10000)		IRAE(×100)		ISE(×10000)	
ψ	$\tau(\psi)$ | estimator	CFG	P	CFG	P	CFG	P	CFG	P	CFG	P	CFG	P
Clayton	1/5	1.44	2.47	2.16	6.62	0.9	1.78	0.84	3.37	0.65	1.31	0.43	1.81
	2/5	1.52	7.84	2.41	83.82	0.93	7.93	0.9	83.3	0.67	7.1	0.46	63.43
	3/5	1.63	34.37	2.87	47.27×10^2	1	36.39	1.06	50.09×10^2	0.7	33.13	0.51	35.04×10^2
	4/5	2.3	52.42×10	5.49	35.15×10^5	1.29	75.85×10	1.73	18.66×10^6	0.86	47.71×10	0.76	26.10×10^5
Frank	1/5	1.38	1.84	2	3.55	0.87	1.16	0.78	1.4	0.62	0.83	0.4	0.72
	2/5	1.37	2.29	1.96	5.52	0.87	1.43	0.77	2.13	0.61	1.03	0.39	1.1
	3/5	1.33	2.89	1.82	9.12	0.83	1.88	0.71	3.67	0.58	1.35	0.36	1.93
	4/5	1.48	4.27	2.26	19.98	0.86	2.81	0.75	8.51	0.58	2.05	0.35	4.43
Gumbel	1/5	1.43	1.8	2.12	3.37	0.93	1.17	0.86	1.41	0.66	0.84	0.46	0.75
	2/5	1.5	2.22	2.37	5.17	0.94	1.46	0.9	2.22	0.69	1.08	0.51	1.21
	3/5	1.59	3.23	2.69	11.29	0.97	2.18	0.98	5	0.72	1.62	0.55	2.74
	4/5	2.01	7.46	4.31	70.59	1.13	5.28	1.3	32.79	0.78	4.47	0.64	23.05
Joe	1/5	1.45	1.7	2.23	3.03	0.94	1.12	0.9	1.28	0.68	0.8	0.48	0.66
	2/5	1.53	1.94	2.52	3.96	0.98	1.29	1.01	1.72	0.73	0.91	0.56	0.86
	3/5	1.74	2.38	3.27	5.82	1.06	1.56	1.19	2.55	0.77	1.11	0.64	1.28
	4/5	2.49	3.27	6.54	11.4	1.33	2.2	1.86	5.02	0.91	1.59	0.87	2.61

Table A.3: Average Integrated relative absolute error (IRAEx100) and Integrated squared error (ISEx10000) of $A_{n,c}^{\mathrm{CFG}}$ and $A_{n,c}^{\mathrm{P}}$ for 4-dimensional Archimax copula $C_{\psi,A}$ samples of size $n \in \{200, 500, 1000\}$. The Pickands dependence function A is **LG** with parameter $\varrho = 2$ so that $\tau(A) = 1/2$, where $\tau(A) = \tau(C_A)$ is the averaged Kendall's tau across all bivariate margins of C_A. There are four choices for the Archimedean generator ψ, Clayton, Frank, Gumbel and Joe, each with four parameter choices so that $\tau(\psi) \in \{1/5, 2/5, 3/5, 4/5\}$, where $\tau(\psi) = \tau(C_\psi)$ is Kendall's tau of the bivariate Archimedean copula C_ψ. There are 1000 Monte Carlo replicates.

	n	200				500				1000			
	error	IRAE(×100)		ISE(×10000)		IRAE(×100)		ISE(×10000)		IRAE(×100)		ISE(×10000)	
ψ	$\tau(\psi)$ | estimator	CFG	P	CFG	P	CFG	P	CFG	P	CFG	P	CFG	P
Clayton	1/5	2.89	4.66	4.23	10.44	1.79	3.29	1.6	5.29	1.29	2.4	0.82	2.82
	2/5	2.91	15.47	4.22	125.2	1.84	14.31	1.71	106.94	1.3	13.99	0.87	98.9
	3/5	3.07	71.13	4.95	67.03×10^2	1.97	69.62	2.01	67.85×10^2	1.41	74.4	1	75.85×10^2
	4/5	4.05	13.12×10^2	8.76	36.57×10^6	2.5	19.46×10^2	3.3	81.52×10^6	1.6	19.42×10^2	1.33	54.47×10^6
Frank	1/5	2.76	3.21	3.86	5.01	1.67	1.99	1.39	1.98	1.23	1.39	0.75	0.96
	2/5	2.75	3.85	3.82	7.23	1.64	2.45	1.37	2.98	1.18	1.72	0.68	1.49
	3/5	2.47	5.3	3.08	14.01	1.54	3.24	1.17	5.26	1.1	2.41	0.6	2.9
	4/5	2.54	8.28	3.37	36.42	1.49	4.88	1.13	11.67	1.05	3.68	0.55	6.76
Gumbel	1/5	2.96	3.17	4.35	4.96	1.82	2.07	1.67	2.15	1.32	1.43	0.86	1.02
	2/5	3.02	3.96	4.57	7.67	1.9	2.66	1.8	3.49	1.34	1.82	0.88	1.66
	3/5	3.13	5.8	5.02	16.48	1.97	4.07	1.96	8.11	1.34	2.88	0.89	4.14
	4/5	3.7	13.8	7.29	10.46×10	2.21	9.94	2.53	50	1.45	7.84	1.06	30.24
Joe	1/5	3.08	3.13	4.7	4.81	1.86	1.99	1.74	1.96	1.37	1.37	0.93	0.94
	2/5	3.23	3.56	5.31	6.28	1.99	2.28	1.96	2.6	1.43	1.61	1.04	1.3
	3/5	3.51	4.34	6.61	9.27	2.17	2.82	2.39	3.93	1.52	1.98	1.18	1.96
	4/5	4.52	6.02	11.26	17.85	2.65	3.88	3.61	7.54	1.78	2.77	1.63	3.86

Table A.4: Average Integrated relative absolute error (IRAEx100) and Integrated squared error (ISEx10000) of $A_{n,c}^{\mathrm{CFG}}$ and $A_{n,c}^{\mathrm{P}}$ for 4-dimensional Archimax copula $C_{\psi,A}$ samples of size $n \in \{200, 500, 1000\}$. The Pickands dependence function A is **NSD** with parameters $\boldsymbol{\alpha} = (1,2,3,4)$, $\rho = 0.59$, $\tau(A) = 1/2$, where $\tau(A) = \tau(C_A)$ is the averaged Kendall's tau across all bivariate margins of C_A. There are four choices for the Archimedean generator ψ, Clayton, Frank, Gumbel and Joe, each with four parameter choices so that $\tau(\psi) \in \{1/5, 2/5, 3/5, 4/5\}$, where $\tau(\psi) = \tau(C_\psi)$ is Kendall's tau of the bivariate Archimedean copula C_ψ. There are 1000 Monte Carlo replicates.

	n	200				500				1000			
	error	IRAE(×100)		ISE(×10000)		IRAE(×100)		ISE(×10000)		IRAE(×100)		ISE(×10000)	
ψ	$\tau(\psi)$ | estimator	CFG	P	CFG	P	CFG	P	CFG	P	CFG	P	CFG	P
Clayton	1/5	2.08	3.53	1.8	5.11	1.34	2.47	0.75	2.56	0.93	1.91	0.36	1.52
	2/5	2.13	12.16	1.91	69.53	1.33	11.54	0.75	63.33	0.94	10.99	0.38	57.14
	3/5	2.41	60.02	2.57	46.99×10^2	1.51	56.85	0.97	45.89×10^2	1.06	55.34	0.47	36.54×10^2
	4/5	3.55	13.92×10^2	5.28	35.28×10^6	1.99	12.48×10^2	1.67	37.26×10^6	1.28	22.95×10^2	0.68	25.54×10^7
Frank	1/5	2.02	2.53	1.65	2.67	1.25	1.6	0.66	1.07	0.88	1.15	0.33	0.55
	2/5	1.9	3.04	1.48	3.82	1.18	2.01	0.6	1.65	0.85	1.37	0.31	0.78
	3/5	1.86	4.07	1.48	6.96	1.19	2.68	0.6	2.98	0.82	1.85	0.28	1.42
	4/5	2.18	6.12	2.02	16.04	1.2	4.03	0.61	6.96	0.77	2.88	0.25	3.45
Gumbel	1/5	2.08	2.5	1.81	2.6	1.34	1.66	0.76	1.14	0.93	1.16	0.36	0.56
	2/5	2.13	3.09	1.93	3.95	1.36	2.07	0.8	1.78	0.96	1.48	0.39	0.91
	3/5	2.31	4.48	2.29	8.34	1.42	3.1	0.87	4.03	0.98	2.27	0.41	2.16
	4/5	3.09	11.23	3.96	61.4	1.7	7.8	1.21	28.63	1.12	6.09	0.53	15.69
Joe	1/5	2.17	2.45	1.97	2.51	1.37	1.6	0.81	1.06	0.96	1.11	0.39	0.51
	2/5	2.28	2.75	2.25	3.11	1.46	1.85	0.92	1.39	1.01	1.27	0.43	0.67
	3/5	2.66	3.43	3.09	4.83	1.6	2.17	1.11	1.94	1.09	1.54	0.51	0.98
	4/5	3.82	4.88	6.06	10.06	2.1	3.02	1.79	3.78	1.36	2.16	0.76	1.92

Table A.5: Average Integrated relative absolute error (IRAEx100) and Integrated squared error (ISEx10000) of $A_{n,c}^{\mathrm{CFG}}$ and $A_{n,c}^{\mathrm{P}}$ for 10-dimensional Archimax copula $C_{\psi,A}$ samples of size $n \in \{200, 500, 1000\}$. The Pickands dependence function A is **LG** with parameter $\varrho = 2$ so that $\tau(A) = 1/2$, where $\tau(A) = \tau(C_A)$ is the averaged Kendall's tau across all bivariate margins of C_A. There are four choices for the Archimedean generator ψ, Clayton, Frank, Gumbel and Joe, each with four parameter choices so that $\tau(\psi) \in \{1/5, 2/5, 3/5, 4/5\}$, where $\tau(\psi) = \tau(C_\psi)$ is Kendall's tau of the bivariate Archimedean copula C_ψ. There are 1000 Monte Carlo replicates.

	n	200				500				1000			
	error	IRAE(×100)		ISE(×10000)		IRAE(×100)		ISE(×10000)		IRAE(×100)		ISE(×10000)	
ψ	$\tau(\psi)$ | estimator	CFG	P	CFG	P	CFG	P	CFG	P	CFG	P	CFG	P
Clayton	1/5	4.76	7.23	4.62	10.27	3.02	4.98	1.89	4.93	2.15	3.89	0.94	3
	2/5	4.98	24.06	5.03	11.74×10	3.1	22.02	2	91.01	2.28	21.08	1.07	82.74
	3/5	5.22	11.19×10	5.58	55.04×10^2	3.3	10.00×10	2.26	44.32×10^2	2.38	10.69×10	1.17	69.58×10^2
	4/5	6.13	85.86×10^2	8.41	31.67×10^8	3.75	33.04×10^2	2.97	11.26×10^7	2.66	59.59×10^2	1.47	48.02×10^7
Frank	1/5	4.46	4.72	4.12	4.56	2.86	2.94	1.69	1.78	2.14	2.03	0.93	0.84
	2/5	4.12	5.56	3.45	6.31	2.76	3.76	1.55	2.86	1.94	2.56	0.77	1.34
	3/5	3.85	8.08	3.02	13.16	2.44	5.03	1.25	5.24	1.8	3.47	0.67	2.48
	4/5	3.51	12.54	2.63	32.74	2.19	7.81	1.01	12.47	1.57	5.55	0.51	6.21
Gumbel	1/5	4.93	4.79	5.05	4.73	3.16	3.12	2.08	2	2.18	2.14	1.01	0.94
	2/5	5.12	5.86	5.5	7.01	3.24	3.89	2.19	3.1	2.21	2.69	1.04	1.5
	3/5	5.32	9.01	6.03	16.06	3.27	6.07	2.24	7.32	2.27	4.48	1.08	3.97
	4/5	5.9	21.96	7.85	10.24×10	3.49	15.54	2.57	48.2	2.38	12.91	1.21	32.38
Joe	1/5	5.09	4.78	5.41	4.77	3.25	3.1	2.21	1.98	2.28	2.16	1.08	0.95
	2/5	5.46	5.66	6.4	6.51	3.42	3.51	2.44	2.55	2.39	2.51	1.2	1.28
	3/5	5.85	6.94	7.47	9.95	3.65	4.35	2.78	3.85	2.53	3.02	1.33	1.85
	4/5	7.31	9.77	12	20.46	4.21	6.22	3.76	7.9	2.8	4.39	1.64	3.91

Table A.6: Average Integrated relative absolute error (IRAEx100) and Integrated squared error (ISEx10000) of $A_{n,c}^{\mathrm{CFG}}$ and $A_{n,c}^{\mathrm{P}}$ for 10-dimensional Archimax copula $C_{\psi,A}$ samples of size $n \in \{200, 500, 1000\}$. The Pickands dependence function A is **NSD** with parameters $\boldsymbol{\alpha} = (1,1,1,1,2,2,2,3,3,4)$, $\rho = 0.69$, $\tau(A) = 1/2$, where $\tau(A) = \tau(C_A)$ is the averaged Kendall's tau across all bivariate margins of C_A. There are four choices for the Archimedean generator ψ, Clayton, Frank, Gumbel and Joe, each with four parameter choices so that $\tau(\psi) \in \{1/5, 2/5, 3/5, 4/5\}$, where $\tau(\psi) = \tau(C_\psi)$ is Kendall's tau of the bivariate Archimedean copula C_ψ. There are 1000 Monte Carlo replicates.

	n	200				500				1000			
	error	IRAE(×100)		ISE(×10000)		IRAE(×100)		ISE(×10000)		IRAE(×100)		ISE(×10000)	
ψ	$\tau(\psi)$ | estimator	CFG	P	CFG	P	CFG	P	CFG	P	CFG	P	CFG	P
Clayton	1/5	4.65	7.27	4.76	10.99	3.05	5.23	1.99	5.59	2.15	3.97	1.02	3.27
	2/5	4.95	24.09	5.19	122.69	3.21	22.63	2.19	106.37	2.18	22.14	1.04	98.17
	3/5	5.28	106	6.13	57.55×10^2	3.43	112.67	2.52	73.89×10^2	2.44	10.85×10	1.33	60.78×10^2
	4/5	6.25	65.85×10^2	8.88	10.76×10^8	3.89	61.49×10^2	3.42	42.84×10^5	2.81	46.92×10^2	1.74	13.06×10^7
Frank	1/5	4.45	4.73	4.28	4.76	2.98	3.03	1.9	1.99	2.13	2.1	0.98	0.98
	2/5	4.24	5.67	3.81	6.82	2.77	3.73	1.68	3.01	1.99	2.52	0.86	1.39
	3/5	3.95	7.63	3.4	12.55	2.55	5.01	1.42	5.38	1.87	3.59	0.78	2.76
	4/5	3.76	12.32	3.15	34.04	2.31	7.9	1.19	13.44	1.68	5.83	0.63	7.36
Gumbel	1/5	4.77	4.76	5.01	4.94	3.09	3.04	2.06	1.99	2.33	2.22	1.18	1.08
	2/5	5.01	6.05	5.49	7.96	3.2	3.93	2.25	3.29	2.34	2.8	1.21	1.69
	3/5	5.33	9.27	6.26	17.84	3.35	6.24	2.47	8.19	2.36	4.48	1.22	4.28
	4/5	6.18	21.9	8.57	103.21	3.65	16.16	3.01	53.95	2.45	12.99	1.36	35.38
Joe	1/5	4.97	4.69	5.54	4.83	3.24	3	2.29	1.98	2.47	2.2	1.34	1.08
	2/5	5.33	5.46	6.49	6.42	3.42	3.54	2.6	2.68	2.6	2.59	1.51	1.45
	3/5	5.82	6.81	7.81	9.88	3.74	4.38	3.12	4.11	2.72	3.17	1.66	2.16
	4/5	7.41	9.42	12.75	19.47	4.32	6.03	4.26	7.93	3.08	4.36	2.12	4.18

Table A.7: Average Integrated relative absolute error (IRAEx100) and Integrated squared error (ISEx10000) of $A_{n,c}^{\mathrm{CFG}}$ for 2-dimensional Archimax copula $C_{\psi,A}$ samples of size $n \in \{200, 500, 1000\}$. The Pickands dependence function A is **LG** with four choices of parameters so that $\tau(A) \in \{1/5, 2/5, 3/5, 4/5\}$, where $\tau(A) = \tau(C_A)$ is Kendall's tau of the bivariate extreme-value copula C_A. There are four choices for the Archimedean generator ψ, Clayton, Frank, Gumbel and Joe, each with four parameter choices so that $\tau(\psi) \in \{1/5, 2/5, 3/5, 4/5\}$, where $\tau(\psi) = \tau(C_\psi)$ is Kendall's tau of the bivariate Archimedean copula C_ψ. There are 1000 Monte Carlo replicates.

		error	IRAE(x100)				ISE(x10000)			
n	ψ	$\tau(\psi)$ | $\tau(A)$	1/5	2/5	3/5	4/5	1/5	2/5	3/5	4/5
200	Clayton	1/5	6.6	3.23	1.03	0.17	2.22	1.72	1.04	0.4
		2/5	8.68	3.9	1.2	0.2	2.54	1.88	1.11	0.44
		3/5	10.01	4.4	1.51	0.36	2.7	1.96	1.23	0.6
		4/5	16.01	7.96	3.49	1.08	3.4	2.62	1.9	1.12
	Frank	1/5	6.49	3.11	1	0.15	2.23	1.7	1.03	0.38
		2/5	7.16	3.19	1.01	0.16	2.31	1.69	1.01	0.39
		3/5	6.94	2.97	0.94	0.18	2.31	1.66	0.99	0.42
		4/5	7.56	3.3	1.24	0.39	2.36	1.72	1.13	0.65
	Gumbel	1/5	7.59	3.46	1.09	0.16	2.39	1.77	1.06	0.4
		2/5	8.95	3.78	1.19	0.19	2.61	1.87	1.12	0.43
		3/5	10.13	4.31	1.39	0.3	2.75	1.96	1.19	0.55
		4/5	13.4	6.17	2.48	0.75	3.1	2.33	1.6	0.93
	Joe	1/5	7.99	3.64	1.14	0.18	2.45	1.82	1.08	0.42
		2/5	9.54	4.09	1.28	0.22	2.7	1.93	1.15	0.47
		3/5	11.67	5.13	1.71	0.37	2.94	2.13	1.31	0.62
		4/5	18.06	8.65	3.74	1.11	3.62	2.75	1.98	1.14
500	Clayton	1/5	2.52	1.23	0.4	0.06	1.38	1.06	0.64	0.24
		2/5	3.25	1.47	0.46	0.07	1.57	1.16	0.69	0.26
		3/5	4.12	1.75	0.55	0.1	1.74	1.24	0.74	0.32
		4/5	5.66	2.58	0.96	0.26	2.0	1.48	0.99	0.55
	Frank	1/5	2.48	1.22	0.38	0.05	1.36	1.05	0.63	0.23
		2/5	2.71	1.21	0.37	0.05	1.44	1.06	0.62	0.23
		3/5	2.78	1.17	0.36	0.06	1.44	1.03	0.61	0.23
		4/5	2.61	1.12	0.38	0.09	1.39	1.0	0.63	0.3
	Gumbel	1/5	3.06	1.42	0.43	0.06	1.52	1.14	0.67	0.24
		2/5	3.7	1.57	0.47	0.07	1.67	1.2	0.69	0.26
		3/5	4.15	1.69	0.51	0.08	1.75	1.22	0.72	0.29
		4/5	4.77	2.06	0.74	0.19	1.88	1.35	0.88	0.46
	Joe	1/5	3.33	1.49	0.45	0.06	1.59	1.16	0.68	0.24
		2/5	4.2	1.74	0.52	0.08	1.77	1.25	0.73	0.27
		3/5	4.94	2.01	0.62	0.11	1.9	1.33	0.78	0.34
		4/5	6.32	2.77	1.03	0.29	2.13	1.56	1.04	0.59
1000	Clayton	1/5	1.42	0.7	0.22	0.03	1.03	0.8	0.48	0.17
		2/5	1.78	0.81	0.25	0.03	1.17	0.86	0.51	0.19
		3/5	2.22	0.95	0.3	0.05	1.3	0.94	0.56	0.22
		4/5	2.89	1.26	0.44	0.11	1.46	1.07	0.68	0.35
	Frank	1/5	1.37	0.68	0.21	0.03	1.02	0.79	0.47	0.16
		2/5	1.5	0.66	0.19	0.03	1.07	0.78	0.45	0.16
		3/5	1.51	0.64	0.19	0.03	1.07	0.76	0.45	0.16
		4/5	1.42	0.6	0.18	0.03	1.04	0.74	0.45	0.19
	Gumbel	1/5	1.66	0.76	0.23	0.03	1.13	0.84	0.49	0.17
		2/5	1.95	0.83	0.25	0.03	1.22	0.87	0.51	0.18
		3/5	2.2	0.9	0.27	0.04	1.3	0.92	0.53	0.2
		4/5	2.54	1.07	0.35	0.07	1.39	1.0	0.61	0.28
	Joe	1/5	1.77	0.78	0.24	0.03	1.16	0.85	0.5	0.18
		2/5	2.17	0.9	0.26	0.04	1.28	0.91	0.53	0.19
		3/5	2.56	1.04	0.31	0.05	1.39	0.98	0.57	0.23
		4/5	3.21	1.38	0.47	0.11	1.56	1.13	0.71	0.36

Table A.8: Average Integrated relative absolute error (IRAEx100) and Integrated squared error (ISEx10000) of $A_{n,c}^{\mathrm{CFG}}$ and for 4-dimensional Archimax copula $C_{\psi,A}$ samples of size $n \in \{200, 500, 1000\}$. The Pickands dependence function A is **LG** with four choices of parameters so that $\tau(A) \in \{1/5, 2/5, 3/5, 4/5\}$, where $\tau(A) = \tau(C_A)$ is Kendall's tau of the corresponding bivariate extreme-value copula $C_{A^{(2)}}$. $A^{(2)}$ is a 2-dimensional **LG** Pickands dependence function with the same parameter as A. There are four choices for the Archimedean generator ψ, Clayton, Frank, Gumbel and Joe, each with four parameter choices so that $\tau(\psi) \in \{1/5, 2/5, 3/5, 4/5\}$, where $\tau(\psi) = \tau(C_\psi)$ is Kendall's tau of the bivariate Archimedean copula C_ψ. There are 1000 Monte Carlo replicates.

		error	IRAE(x100)				ISE(x10000)			
n	ψ	$\tau(\psi) \mid \tau(A)$	1/5	2/5	3/5	4/5	1/5	2/5	3/5	4/5
$n = 200$	Clayton	1/5	17.84	7.55	2.05	0.26	4.33	3.49	2.22	0.9
		2/5	20.5	7.71	2.08	0.3	4.65	3.56	2.23	0.96
		3/5	25.09	8.98	2.58	0.52	5.12	3.76	2.42	1.28
		4/5	37.15	14.51	5.2	1.58	5.92	4.68	3.49	2.4
	Frank	1/5	17.12	7.02	1.88	0.24	4.24	3.36	2.11	0.85
		2/5	18.62	7.03	1.83	0.24	4.45	3.38	2.09	0.86
		3/5	16.1	5.68	1.52	0.27	4.06	3.02	1.9	0.92
		4/5	15.33	5.72	1.97	0.59	3.93	3	2.15	1.46
	Gumbel	1/5	20.93	8.01	2.06	0.25	4.71	3.62	2.24	0.9
		2/5	24.64	8.63	2.21	0.32	5.04	3.74	2.3	1
		3/5	27.75	9.35	2.5	0.46	5.3	3.84	2.43	1.22
		4/5	33.36	12.27	4.17	1.14	5.71	4.32	3.12	2.05
	Joe	1/5	23.9	8.77	2.23	0.28	5.03	3.79	2.33	0.94
		2/5	29.48	10.16	2.53	0.36	5.49	4.02	2.45	1.06
		3/5	35.24	12.21	3.31	0.63	5.85	4.32	2.74	1.42
		4/5	51.43	19.1	6.54	1.74	6.81	5.25	3.87	2.54
$n = 500$	Clayton	1/5	6.95	2.89	0.77	0.09	2.71	2.17	1.37	0.54
		2/5	7.63	3.11	0.84	0.11	2.8	2.24	1.41	0.6
		3/5	9.81	3.67	1.01	0.17	3.15	2.41	1.54	0.74
		4/5	14.66	5.6	1.88	0.45	3.78	2.93	2.1	1.31
	Frank	1/5	6.15	2.54	0.67	0.08	2.53	2.03	1.27	0.51
		2/5	6.24	2.47	0.66	0.08	2.53	1.99	1.26	0.51
		3/5	6.08	2.17	0.57	0.08	2.52	1.89	1.18	0.52
		4/5	5.74	2.02	0.6	0.13	2.43	1.8	1.19	0.67
	Gumbel	1/5	8.01	3.06	0.8	0.1	2.9	2.23	1.39	0.55
		2/5	9.74	3.38	0.86	0.11	3.19	2.34	1.44	0.6
		3/5	11.18	3.67	0.96	0.14	3.41	2.43	1.51	0.68
		4/5	13.1	4.49	1.35	0.31	3.62	2.65	1.78	1.07
	Joe	1/5	8.95	3.26	0.83	0.1	3.06	2.3	1.42	0.56
		2/5	11.12	3.71	0.94	0.12	3.39	2.46	1.51	0.63
		3/5	13.39	4.5	1.16	0.19	3.66	2.67	1.67	0.79
		4/5	17.54	6.34	1.96	0.48	4.15	3.13	2.18	1.34
$n = 1000$	Clayton	1/5	3.47	1.48	0.4	0.05	1.9	1.56	0.99	0.39
		2/5	4.01	1.59	0.42	0.05	2.01	1.58	1	0.41
		3/5	5.11	1.85	0.48	0.07	2.3	1.73	1.08	0.48
		4/5	6.46	2.35	0.71	0.16	2.52	1.9	1.3	0.76
	Frank	1/5	3.22	1.36	0.36	0.04	1.81	1.48	0.94	0.37
		2/5	3.26	1.26	0.32	0.04	1.86	1.45	0.89	0.35
		3/5	3.23	1.14	0.28	0.04	1.83	1.36	0.83	0.34
		4/5	2.9	1.02	0.27	0.05	1.73	1.28	0.82	0.4
	Gumbel	1/5	4.31	1.61	0.41	0.05	2.12	1.62	1	0.39
		2/5	4.91	1.67	0.41	0.05	2.27	1.66	1.01	0.4
		3/5	5.27	1.73	0.42	0.06	2.34	1.68	1.02	0.43
		4/5	5.86	1.97	0.54	0.11	2.45	1.77	1.14	0.61
	Joe	1/5	4.87	1.74	0.43	0.05	2.25	1.69	1.03	0.4
		2/5	6.08	2	0.48	0.06	2.49	1.79	1.08	0.43
		3/5	7.1	2.31	0.56	0.08	2.66	1.9	1.15	0.5
		4/5	8.69	2.97	0.85	0.18	2.91	2.15	1.43	0.8

Table A.9: Average Integrated relative absolute error (IRAEx100) and Integrated squared error (ISEx10000) of $A_{n,c}^{\mathrm{CFG}}$ and for 10-dimensional Archimax copula $C_{\psi,A}$ samples of size $n \in \{200, 500, 1000\}$. The Pickands dependence function A is **LG** with four choices of parameters so that $\tau(A) \in \{1/5, 2/5, 3/5, 4/5\}$, where $\tau(A) = \tau(C_A)$ is Kendall's tau of the corresponding bivariate extreme-value copula $C_{A^{(2)}}$. $A^{(2)}$ is a 2-dimensional **LG** Pickands dependence function with the same parameter as A. There are four choices for the Archimedean generator ψ, Clayton, Frank, Gumbel and Joe, each with four parameter choices so that $\tau(\psi) \in \{1/5, 2/5, 3/5, 4/5\}$, where $\tau(\psi) = \tau(C_\psi)$ is Kendall's tau of the bivariate Archimedean copula C_ψ. There are 1000 Monte Carlo replicates.

		error	IRAE(x100)				ISE(x10000)			
n	ψ	$\tau(\psi) \mid \tau(A)$	1/5	2/5	3/5	4/5	1/5	2/5	3/5	4/5
$n = 200$	Clayton	1/5	32.88	9.78	1.97	0.2	7.1	5.74	3.73	1.6
		2/5	35.12	9.92	2.03	0.23	7.36	5.7	3.7	1.72
		3/5	43.17	11.68	2.48	0.38	8.1	6.22	4.1	2.19
		4/5	54.62	15.64	4.19	1.08	8.87	7.01	5.21	3.93
	Frank	1/5	31.49	9.01	1.82	0.18	7	5.5	3.54	1.55
		2/5	28.42	7.96	1.62	0.18	6.61	5.15	3.34	1.51
		3/5	26.69	6.88	1.36	0.19	6.43	4.82	3.07	1.57
		4/5	21.7	5.79	1.46	0.4	5.67	4.3	3.1	2.39
	Gumbel	1/5	40.74	10.81	2.05	0.21	7.89	6.02	3.77	1.64
		2/5	48.43	11.84	2.18	0.24	8.54	6.31	3.91	1.76
		3/5	52.22	12.45	2.36	0.35	8.74	6.37	3.99	2.12
		4/5	61.23	15.44	3.7	0.85	9.08	6.79	4.88	3.47
	Joe	1/5	47.39	11.98	2.23	0.22	8.46	6.3	3.91	1.68
		2/5	59.62	14.23	2.59	0.29	9.27	6.76	4.16	1.89
		3/5	71.83	16.91	3.2	0.45	9.86	7.13	4.51	2.39
		4/5	95.53	24.21	5.59	1.19	10.97	8.3	5.97	4.15
$n = 500$	Clayton	1/5	13.8	4.34	0.87	0.08	4.59	3.82	2.47	1.04
		2/5	16.11	4.59	0.89	0.09	4.94	3.93	2.5	1.08
		3/5	18.48	5.05	1.01	0.12	5.22	4.07	2.63	1.28
		4/5	23.39	6.41	1.52	0.31	5.91	4.57	3.21	2.1
	Frank	1/5	13.16	4.01	0.78	0.07	4.48	3.66	2.32	0.97
		2/5	11.96	3.41	0.67	0.07	4.24	3.34	2.14	0.92
		3/5	10.51	2.74	0.53	0.06	3.95	3	1.91	0.89
		4/5	8.29	2.13	0.47	0.09	3.47	2.62	1.78	1.14
	Gumbel	1/5	17.62	4.68	0.88	0.08	5.1	3.93	2.47	1.04
		2/5	21.52	5.16	0.93	0.09	5.61	4.12	2.54	1.09
		3/5	23.46	5.45	0.98	0.11	5.85	4.22	2.61	1.22
		4/5	24.49	5.89	1.25	0.22	5.96	4.35	2.9	1.78
	Joe	1/5	20.08	5.01	0.93	0.09	5.45	4.05	2.54	1.08
		2/5	25.61	5.85	1.04	0.1	6.14	4.41	2.71	1.18
		3/5	29.34	6.66	1.19	0.14	6.51	4.67	2.87	1.38
		4/5	34.69	8.39	1.81	0.34	6.93	5.1	3.47	2.2
$n = 1000$	Clayton	1/5	6.53	2.03	0.4	0.04	3.11	2.6	1.68	0.71
		2/5	7.1	2.09	0.42	0.04	3.21	2.62	1.71	0.74
		3/5	8.9	2.38	0.47	0.05	3.65	2.81	1.8	0.83
		4/5	10.7	2.76	0.6	0.11	3.99	3.01	2.03	1.23
	Frank	1/5	6.57	2.01	0.39	0.04	3.11	2.59	1.65	0.69
		2/5	5.9	1.65	0.32	0.03	2.97	2.35	1.5	0.63
		3/5	5.4	1.42	0.27	0.03	2.83	2.17	1.38	0.61
		4/5	4.06	1.05	0.22	0.03	2.43	1.85	1.22	0.67
	Gumbel	1/5	8.53	2.23	0.41	0.04	3.56	2.72	1.7	0.7
		2/5	9.91	2.38	0.43	0.04	3.81	2.8	1.74	0.73
		3/5	10.61	2.49	0.45	0.05	3.94	2.86	1.77	0.79
		4/5	11.09	2.68	0.54	0.08	4.04	2.99	1.95	1.09
	Joe	1/5	10.08	2.47	0.45	0.04	3.85	2.85	1.76	0.73
		2/5	12.37	2.82	0.49	0.05	4.18	3.01	1.84	0.79
		3/5	13.83	3.13	0.56	0.06	4.42	3.16	1.95	0.9
		4/5	15.34	3.7	0.76	0.13	4.64	3.42	2.28	1.35

Table A.10: Average Integrated relative absolute error (IRAEx100) and Integrated squared error (ISEx10000) of $A_{n,c}^{\mathrm{CFG}}$ and for d-dimensional Archimax copula $C_{\psi,A}$ samples of size $n \in \{200, 500, 1000\}$ for dimensions $d \in \{2, 4, 10\}$. The Pickands dependence function A is **LG** with four choices of parameters so that $\lambda_U(A) \in \{1/5, 2/5, 3/5, 4/5\}$, where $\lambda_U(A) = \lambda_U(C_A)$ is the upper tail dependence coefficient of the corresponding bivariate extreme-value copula $C_{A^{(2)}}$. $A^{(2)}$ is a 2-dimensional **LG** Pickands dependence function with the same parameter as A. The Archimedean generator ψ is Joe and $\lambda_U(\psi, A) \in \{5/10, 6/10, 7/10, 8/10, 9/10\}$, where $\lambda_U(\psi, A) = \lambda_U(C_{\psi,A})$ is the upper tail coefficient of the bivariate Archimax copula $C_{\psi,A^{(2)}}$. There are 1000 Monte Carlo replicates.

		error	IRAE(x100)				ISE(x10000)			
	d	$\lambda_U(\psi, A) \mid \lambda_U(A)$	1/5	2/5	3/5	4/5	1/5	2/5	3/5	4/5
$n = 200$	2	5/10	9.34	4.63			2.6	1.98		
		6/10	10.37	5.04			2.74	2.07		
		7/10	11.38	5.49	1.94		2.88	2.16	1.38	
		8/10	13.18	6.11	2.11		3.07	2.26	1.45	
		9/10	17.3	8.22	2.73	.39	3.48	2.59	1.61	.63
	4	5/10	29.13	11.76			5.25	4.02		
		6/10	33.02	13.18			5.55	4.27		
		7/10	37.03	14.6	4		5.82	4.47	2.89	
		8/10	41.83	16.44	4.47		6.08	4.66	3.03	
		9/10	54.45	20.81	5.78	.65	6.73	5.1	3.34	1.39
	10	5/10	63.58	17.87			8.88	6.64		
		6/10	73.1	20.78			9.42	7.14		
		7/10	82.19	23.37	4.53		9.89	7.51	4.78	
		8/10	94.21	27.01	5.21		10.34	7.87	5.12	
		9/10	112.8	31.95	6.53	.56	10.93	8.37	5.52	2.44
$n = 500$	2	5/10	3.91	1.83			1.67	1.26		
		6/10	4.45	2.08			1.78	1.33		
		7/10	5.02	2.35	.77		1.89	1.41	.88	
		8/10	5.64	2.63	.88		1.98	1.48	.93	
		9/10	6.65	3.11	1.03	.14	2.13	1.59	.99	.38
	4	5/10	10.98	4.37			3.2	2.44		
		6/10	12.59	4.88			3.42	2.59		
		7/10	14.08	5.45	1.52		3.6	2.74	1.76	
		8/10	16.22	6.12	1.7		3.83	2.9	1.88	
		9/10	19.48	7.63	2.12	.24	4.16	3.17	2.07	.86
	10	5/10	27.19	7.63			5.74	4.32		
		6/10	31.54	8.71			6.16	4.58		
		7/10	35.6	9.87	1.9		6.55	4.87	3.11	
		8/10	39.37	11.02	2.17		6.88	5.17	3.34	
		9/10	44.77	12.64	2.52	.22	7.19	5.46	3.58	1.58
$n = 1000$	2	5/10	2.06	1			1.22	.92		ÊÊ
		6/10	2.31	1.1			1.3	.98		
		7/10	2.6	1.22	.41		1.37	1.02	.64	
		8/10	2.95	1.37	.45		1.45	1.08	.67	
		9/10	3.47	1.62	.53	.07	1.58	1.17	.73	.28
	4	5/10	5.98	2.38			2.36	1.81		
		6/10	6.8	2.64			2.5	1.9		
		7/10	7.71	2.95	.8		2.65	2.01	1.29	
		8/10	8.68	3.31	.89		2.79	2.1	1.36	
		9/10	10.12	3.88	1.05	.12	2.97	2.25	1.45	.6
	10	5/10	13.72	3.82			4.06	3.07		
		6/10	15.62	4.31			4.28	3.24		
		7/10	17.26	4.84	.93		4.47	3.39	2.19	
		8/10	18.7	5.33	1.04		4.65	3.53	2.29	
		9/10	20.29	5.89	1.2	.11	4.85	3.7	2.44	1.08

Table A.11: Average Integrated relative absolute error (IRAEx100) and Integrated squared error (ISEx10000) of $A_{n,c}^{\mathrm{CFG}}$ and $A_{n,c}^{\mathrm{P}}$ for d-dimensional Archimax copula $C_{\psi,A}$ samples of size $n \in \{200, 500, 1000\}$ for dimensions $d \in \{2, 4, 10\}$. The Pickands dependence function A is **LG** with four choices of parameters so that $\eta_L(A) \in \{0.57, 0.66, 0.76, 0.87\}$, where $\eta_L(A) = \eta_L(C_A)$ is the index of lower tail dependence Ledford and Tawn (1996) of the corresponding bivariate extreme-value copula $C_{A^{(2)}}$. $A^{(2)}$ is a 2-dimensional **LG** Pickands dependence function with the same parameter as A. The Archimedean generator ψ is Clayton and $\lambda_L(\psi, A) \in \{1/5, 2/5, 3/5, 4/5\}$, where $\lambda_L(\psi, A) = \lambda_L(C_{\psi,A})$ is the lower tail coefficient of the bivariate Archimax copula $C_{\psi,A^{(2)}}$. There are 1000 Monte Carlo replicates.

		error	IRAE(x100)				ISE(x10000)			
	d	$\lambda_L(\psi, A) \mid \eta_L(A)$	0.57	0.66	0.76	0.87	0.57	0.66	0.76	0.87
$n = 200$	2	1/5	6.39	3.24	1.06	0.15	2.17	1.71	1.05	0.39
		2/5	6.81	3.23	1.04	0.16	2.26	1.71	1.04	0.39
		3/5	8.33	3.47	1.07	0.15	2.54	1.79	1.06	0.39
		4/5	9.96	4.07	1.22	0.17	2.71	1.9	1.12	0.4
	4	1/5	16.81	6.99	1.92	0.23	4.22	3.4	2.16	0.86
		2/5	18.08	7.21	2	0.23	4.37	3.42	2.18	0.85
		3/5	20.02	7.42	2.07	0.25	4.58	3.46	2.22	0.89
		4/5	24.27	8.11	2.13	0.27	4.97	3.6	2.26	0.91
	10	1/5	33.23	10.06	2	0.19	7.15	5.8	3.72	1.58
		2/5	34.7	10.12	2	0.2	7.37	5.86	3.74	1.61
		3/5	34.72	10.49	2.02	0.19	7.32	5.94	3.78	1.59
		4/5	39.04	10.01	1.94	0.21	7.68	5.8	3.7	1.63
$n = 500$	2	1/5	2.48	1.26	0.4	0.05	1.36	1.07	0.64	0.23
		2/5	2.78	1.31	0.39	0.05	1.46	1.09	0.65	0.23
		3/5	3.24	1.48	0.4	0.06	1.57	1.16	0.65	0.24
		4/5	3.79	1.5	0.46	0.06	1.69	1.17	0.69	0.24
	4	1/5	6.52	2.88	0.77	0.09	2.62	2.15	1.36	0.53
		2/5	6.73	2.89	0.78	0.09	2.64	2.17	1.37	0.54
		3/5	7.9	2.95	0.75	0.09	2.84	2.18	1.35	0.53
		4/5	9.22	3.21	0.84	0.09	3.1	2.29	1.41	0.55
	10	1/5	12.8	4.15	0.83	0.08	4.4	3.7	2.4	1.0
		2/5	13.81	4.17	0.79	0.08	4.57	3.77	2.35	1.01
		3/5	15.36	4.09	0.86	0.08	4.81	3.71	2.46	1.0
		4/5	17.48	4.89	0.93	0.08	5.1	4.05	2.56	1.05
$n = 1000$	2	1/5	1.37	0.69	0.22	0.03	1.02	0.8	0.48	0.17
		2/5	1.52	0.68	0.22	0.03	1.07	0.8	0.48	0.17
		3/5	1.71	0.73	0.23	0.03	1.14	0.81	0.49	0.17
		4/5	2.07	0.84	0.24	0.03	1.26	0.88	0.5	0.17
	4	1/5	3.41	1.45	0.39	0.04	1.88	1.53	0.97	0.37
		2/5	3.5	1.5	0.38	0.04	1.9	1.56	0.96	0.38
		3/5	3.73	1.46	0.39	0.04	1.96	1.53	0.97	0.38
		4/5	4.84	1.66	0.43	0.05	2.24	1.64	1.01	0.39
	10	1/5	6.54	2.02	0.39	0.04	3.09	2.6	1.63	0.7
		2/5	6.68	1.97	0.4	0.04	3.12	2.56	1.68	0.69
		3/5	7.11	2.08	0.42	0.04	3.22	2.62	1.71	0.71
		4/5	8.37	2.17	0.42	0.04	3.49	2.67	1.72	0.7

Table A.12: Average Integrated relative absolute error (IRAEx100) and Integrated squared error (ISEx10000) of $A_{n,c}^{\mathrm{CFG}}$ for 2-dimensional Archimax copula $C_{\psi,A}$ samples of size $n \in \{200, 500, 1000\}$. The Pickands dependence function A is either **DSM** or **LG**. Parameters for the **DSM** case are reported in Equation (A.1). The parameter in **LG** is set $\varrho = 2.87$ so that the averaged pairwise Kendall's tau of both Pickands dependence functions is approximately equal to 0.65. There are four choices for the Archimedean generator ψ, Clayton, Frank, Gumbel and Joe, each with four parameter choices so that $\tau(\psi) \in \{1/5, 2/5, 3/5, 4/5\}$, where $\tau(\psi) = \tau(C_\psi)$ is Kendall's tau of the bivariate Archimedean copula C_ψ. There are 1000 Monte Carlo replicates.

	error	IRAE(x100)						ISE(x10000)					
	n	200		500		1000		200		500		1000	
ψ	$\tau(\psi)$ | A	DSM	LG	DSM	LG	DSM	LG	DSM	LG	DSM	LG	DSM	LG
Clayton	1/5	0.99	0.9	0.61	0.55	0.42	0.4	1	0.75	0.39	0.28	0.18	0.14
	2/5	1.02	0.96	0.64	0.59	0.44	0.41	1.05	0.87	0.41	0.33	0.19	0.16
	3/5	1.16	1.09	0.71	0.65	0.47	0.45	1.33	1.18	0.5	0.41	0.22	0.19
	4/5	1.76	1.66	0.94	0.92	0.61	0.56	2.89	2.56	0.82	0.78	0.34	0.29
Frank	1/5	0.95	0.88	0.58	0.56	0.41	0.38	0.93	0.73	0.35	0.29	0.17	0.14
	2/5	0.94	0.86	0.57	0.54	0.4	0.38	0.9	0.7	0.33	0.28	0.16	0.13
	3/5	0.93	0.88	0.57	0.54	0.38	0.37	0.88	0.74	0.33	0.27	0.15	0.13
	4/5	1.14	1	0.62	0.56	0.39	0.37	1.22	0.94	0.38	0.29	0.15	0.13
Gumbel	1/5	1	0.91	0.61	0.57	0.41	0.39	1.04	0.78	0.38	0.31	0.17	0.15
	2/5	1.03	0.92	0.63	0.59	0.42	0.41	1.07	0.82	0.41	0.32	0.18	0.16
	3/5	1.12	1.01	0.67	0.63	0.45	0.43	1.23	0.99	0.44	0.37	0.19	0.17
	4/5	1.56	1.38	0.84	0.8	0.52	0.5	2.21	1.77	0.66	0.59	0.26	0.23
Joe	1/5	1.02	0.92	0.64	0.57	0.41	0.41	1.09	0.81	0.41	0.31	0.17	0.16
	2/5	1.12	1	0.68	0.61	0.44	0.44	1.29	0.95	0.46	0.35	0.2	0.18
	3/5	1.27	1.18	0.76	0.66	0.48	0.47	1.63	1.35	0.56	0.42	0.23	0.21
	4/5	1.91	1.81	1.03	0.94	0.61	0.6	3.43	2.98	0.98	0.8	0.35	0.34

Table A.13: Average Integrated relative absolute error (IRAEx100) and Integrated squared error (ISEx10000) of $A_{n,c}^{\mathrm{CFG}}$ for 4-dimensional Archimax copula $C_{\psi,A}$ samples of size $n \in \{200, 500, 1000\}$. The Pickands dependence function A is either **DSM** or **LG**. Parameters for the **DSM** case are reported in Equation (A.2). The parameter in **LG** is set $\varrho = 2.17$ so that the averaged pairwise Kendall's tau of both Pickands dependence functions is approximately equal to 0.54. There are four choices for the Archimedean generator ψ, Clayton, Frank, Gumbel and Joe, each with four parameter choices so that $\tau(\psi) \in \{1/5, 2/5, 3/5, 4/5\}$, where $\tau(\psi) = \tau(C_\psi)$ is Kendall's tau of the bivariate Archimedean copula C_ψ. There are 1000 Monte Carlo replicates.

	error	IRAE(x100)						ISE(x10000)					
	n	200		500		1000		200		500		1000	
ψ	$\tau(\psi) \mid A$	DSM	LG	DSM	LG	DSM	LG	DSM	LG	DSM	LG	DSM	LG
Clayton	1/5	3.59	2.54	2.31	1.63	1.63	1.16	10.98	2.92	4.6	1.22	2.26	0.63
	2/5	4.06	2.63	2.57	1.69	1.9	1.16	14.02	3.17	5.77	1.31	3.08	0.63
	3/5	4.68	2.86	2.98	1.81	2.15	1.28	18.66	3.86	7.67	1.54	3.98	0.77
	4/5	5.89	3.84	3.69	2.28	2.57	1.46	30.67	7.22	11.87	2.48	5.83	0.99
Frank	1/5	3.5	2.47	2.26	1.61	1.61	1.08	10.3	2.81	4.39	1.19	2.25	0.54
	2/5	3.73	2.35	2.4	1.49	1.71	1.07	11.77	2.55	4.96	1.02	2.51	0.53
	3/5	3.9	2.2	2.45	1.42	1.8	1	12.81	2.23	5.2	0.94	2.79	0.45
	4/5	3.97	2.37	2.41	1.42	1.74	0.92	13.73	2.67	5.05	0.94	2.61	0.39
Gumbel	1/5	4.04	2.62	2.57	1.64	1.83	1.18	14.17	3.14	5.6	1.25	2.83	0.64
	2/5	4.52	2.68	2.9	1.68	2.04	1.17	18.11	3.4	7.16	1.31	3.55	0.64
	3/5	4.89	2.82	3.13	1.75	2.2	1.18	21.7	3.86	8.34	1.44	4.14	0.65
	4/5	5.61	3.42	3.38	1.99	2.36	1.28	28.32	5.71	10.1	1.92	4.84	0.78
Joe	1/5	4.39	2.62	2.72	1.66	1.94	1.16	16.65	3.23	6.29	1.27	3.24	0.64
	2/5	5.12	2.86	3.14	1.75	2.19	1.25	23.51	3.94	8.43	1.45	4.17	0.74
	3/5	5.61	3.27	3.43	1.91	2.39	1.34	29.86	5.36	10.25	1.74	5.02	0.86
	4/5	6.88	4.44	3.92	2.45	2.67	1.59	46.17	9.75	13.7	2.83	6.35	1.2

Table A.14: Average Integrated relative absolute error (IRAEx100) and Integrated squared error (ISEx10000) of $A_{n,c}^{\mathrm{CFG}}$ for 10-dimensional Archimax copula $C_{\psi,A}$ samples of size $n \in \{200, 500, 1000\}$. The Pickands dependence function A is either **DSM** or **LG**. Parameters for the **DSM** case are reported in Equation (A.3). The parameter in **LG** is set $\varrho = 1.56$ so that the averaged pairwise Kendall's tau of both Pickands dependence functions is approximately equal to 0.36. There are four choices for the Archimedean generator ψ, Clayton, Frank, Gumbel and Joe, each with four parameter choices so that $\tau(\psi) \in \{1/5, 2/5, 3/5, 4/5\}$, where $\tau(\psi) = \tau(C_\psi)$ is Kendall's tau of the bivariate Archimedean copula C_ψ. There are 1000 Monte Carlo replicates.

	error	IRAE(x100)						ISE(x10000)					
	n	200		500		1000		200		500		1000	
ψ	$\tau(\psi)$ | A	DSM	LG	DSM	LG	DSM	LG	DSM	LG	DSM	LG	DSM	LG
Clayton	1/5	4.97	6.29	2.95	3.94	1.98	2.8	15.32	13.61	5.53	5.5	2.52	2.79
	2/5	5.95	6.21	3.72	4.19	2.55	2.87	22.12	13.15	8.87	6.12	4.23	3
	3/5	6.98	6.73	4.57	4.39	3.24	3.06	30.76	16.11	13.55	6.84	6.92	3.36
	4/5	8.71	7.74	5.47	4.66	3.92	3.45	51.59	23.07	19.81	8.14	10.37	4.37
Frank	1/5	4.84	5.93	2.92	3.75	1.96	2.64	14.75	12.45	5.4	4.98	2.52	2.46
	2/5	5.32	5.65	3.27	3.66	2.25	2.58	17.93	10.99	6.84	4.68	3.28	2.35
	3/5	5.38	5.15	3.35	3.33	2.42	2.34	18.69	9.37	7.3	3.89	3.81	1.94
	4/5	5.27	4.5	3.31	2.91	2.36	2.01	18.6	7.57	7.18	2.97	3.63	1.47
Gumbel	1/5	6.15	6.4	3.99	4.15	2.8	3	23.97	14.72	10.39	6	5.09	3.15
	2/5	7.22	6.61	4.61	4.1	3.37	3.05	34.41	16.19	13.98	6.08	7.32	3.32
	3/5	7.92	6.91	4.89	4.11	3.65	3.11	43.01	17.88	15.75	6.17	8.64	3.51
	4/5	8.84	7.55	5.28	4.33	3.89	3.13	55.56	21.93	18.64	6.99	9.77	3.58
Joe	1/5	6.91	6.84	4.51	4.49	3.23	3.11	30.72	17	13.43	7.17	6.75	3.5
	2/5	8.18	7.33	5.29	4.71	3.9	3.26	45.88	19.59	18.81	8.04	9.84	3.86
	3/5	9.12	7.74	5.74	4.96	4.24	3.37	58.16	22.54	22.55	8.99	11.89	4.13
	4/5	10.85	8.88	6.45	5.44	4.6	3.6	84.1	31.69	28.98	11.25	14.08	4.64

List of Figures

List of Tables

List of publications

Simon Chatelain, Anne-Laure Fougères & Johanna G. Nešlehová (2019). Inference for Archimax copulas. *The Annals of Statistics*, 47, in press.

Simon Chatelain, Anne-Laure Fougères & Johanna G. Nešlehová. Clustered Archimax copulas. *(to be submitted soon)*

Simon Chatelain, Samuel Perreault. Modeling clustered subasymptotic extremes. *(in preparation)*

Bibliography

Abramowitz, M. and Stegun, I. A. (1964). *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. National Bureau of Standards, Washington, DC.

Aulbach, S., Falk, M., and Zott, M. (2015). The space of D-norms revisited. *Extremes*, 18(1):85–97.

Bacro, J.-N., Gaetan, C., Opitz, T., and Toulemonde, G. (2019). Hierarchical space-time modeling of asymptotically independent exceedances with an application to precipitation data. *Journal of the American Statistical Association*, pages 1–33.

Barbe, P., Genest, C., Ghoudi, K., and Rémillard, B. (1996). On Kendall's process. *J. Multivariate Anal.*, 58:197–229.

Beirlant, J., Goegebeur, U., Segers, J., and Teugels, J. L. (2004). *Statistics of Extremes: Theory and Applications*. Wiley.

Belzile, L. R. and Nešlehová, J. G. (2017). Extremal attractors of liouville copulas. *J. Multivariate Anal.*, 160(Supplement C):68–92.

Ben Ghorbal, N., Genest, C., and Nešlehová, J. (2009). On the Ghoudi, Khoudraji, and Rivest test for extreme-value dependence. *Canad. J. Statist.*, 37:534–552.

Berghaus, B., Bücher, A., and Volgushev, S. (2017). Weak convergence of the empirical copula process with respect to weighted metrics. *Bernoulli*, 23:743–772.

Bernard, E., Naveau, P., Vrac, M., and Mestre, O. (2013). Clustering of maxima: Spatial dependencies among heavy rainfall in france. *Journal of Climate*, 26(20):7929–7937.

Bingham, N. H., Goldie, C. M., and Teugels, J. L. (1989). *Regular variation*, volume 27 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge.

Breiman, L. (1965). On some limit theorems similar to the arc-sin law. *Teor. Verojatnost. i Primenen.*, 10:351–360.

Bücher, A. and Ruppert, M. (2013). Consistent testing for a constant copula under strong mixing based on the tapered block multiplier technique. *J. Multivariate Anal.*, 116:208–229.

Bücher, A., Segers, J., and Volgushev, S. (2014). When uniform weak convergence fails: empirical processes for dependence functions and residuals via epi- and hypographs. *Ann. Statist.*, 42(4):1598–1634.

Bücher, A. and Volgushev, S. (2013). Empirical and sequential empirical copula processes under serial dependence. *J. Multivariate Anal.*, 119:61–70.

Bücher, A., Volgushev, S., and Zou, N. (2019). On second order conditions in the multivariate block maxima and peak over threshold method. *Journal of Multivariate Analysis.*

Capéraà, P., Fougères, A.-L., and Genest, C. (1997). A nonparametric estimation procedure for bivariate extreme value copulas. *Biometrika*, 84:567–577.

Capéraà, P., Fougères, A.-L., and Genest, C. (2000). Bivariate distributions with given extreme value attractor. *J. Multivariate Anal.*, 72(1):30–49.

Charpentier, A., Fougères, A.-L., Genest, C., and Nešlehová, J. G. (2014). Multivariate Archimax copulas. *J. Multivariate Anal.*, 126:118–136.

Charpentier, A. and Segers, J. (2009). Tails of multivariate Archimedean copulas. *J. Multivariate Anal.*, 100:1521–1537.

Cherubini, U., Luciano, E., and Vecchiato, W. (2004). *Copula methods in finance.* John Wiley & Sons.

Coles, S. (2001). *An introduction to statistical modeling of extreme values.* Springer Series in Statistics. Springer-Verlag London, Ltd., London.

Coles, S., Heffernan, J., and Tawn, J. (1999). Dependence measures for extreme value analyses. *Extremes*, 2:339–365.

Cooley, D., Naveau, P., and Poncet, P. (2006). Variograms for spatial max-stable random fields. In *Dependence in probability and statistics*, pages 373–390. Springer.

De Haan, L. et al. (1984). A spectral representation for max-stable processes. *The annals of probability*, 12(4):1194–1204.

de Haan, L. and Ferreira, A. (2006). *Extreme Value Theory.* Springer, New York.

de Haan, L. and Resnick, S. I. (1977). Limit theory for multivariate sample extremes. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, 40(4):317–337.

Deheuvels, P. (1979). La fonction de dépendance empirique et ses propriétés. Un test non paramétrique d'indépendance. *Acad. Roy. Belg. Bull. Cl. Sci. (5)*, 65(6):274–292.

Deheuvels, P. (1981a). An asymptotic decomposition for multivariate distribution-free tests of independence. *J. Multivariate Anal.*, 11(1):102–113.

Deheuvels, P. (1981b). A Kolmogorov-Smirnov type test for independence and multivariate samples. *Rev. Roumaine Math. Pures Appl.*, 26(2):213–226.

Di Bernardino, E. and Rulliere, D. (2013). On certain transformations of archimedean copulas: Application to the non-parametric estimation of their generators. *Dependence Modeling*, 1:1–36.

Einmahl, J., Kiriliouk, A., and Segers, J. (2017). A continuous updating weighted least squares estimator of tail dependence in high dimension. *Extremes*, 21:1–29.

Embrechts, P. and Goldie, C. M. (1980). On closure and factorization properties of subexponential and related distributions. *J. Austral. Math. Soc. Ser. A*, 29(2):243–256.

Embrechts, P., Klüppelberg, C., and Mikosch, T. (1997). *Modelling Extremal Events for Insurance and Finance*. Springer, New York.

Embrechts, P., McNeil, A. J., and Straumann, D. (2002). Correlation and dependence in risk management: properties and pitfalls. In *Risk management: value at risk and beyond (Cambridge, 1998)*, pages 176–223. Cambridge Univ. Press, Cambridge.

Engelke, S., Opitz, T., and Wadsworth, J. (2019). Extremal dependence of random scale constructions. *arXiv preprint arXiv:1803.04221*.

Fermanian, J.-D., Radulović, D., and Wegkamp, M. (2004). Weak convergence of empirical copula processes. *Bernoulli*, 10(5):847–860.

Fisher, R. A. and Tippett, L. H. C. (1928). Limiting forms of the frequency distribution of the largest or smallest member of a sample. In *Mathematical Proceedings of the Cambridge Philosophical Society*, volume 24, pages 180–190. Cambridge University Press.

Fougères, A.-L., de Haan, L., and Mercadier, C. (2015). Bias correction in multivariate extremes. *Ann. Statist.*, 43:903–934.

Fougères, A.-L., Mercadier, C., and Nolan, J. P. (2013). Dense classes of multivariate extreme value distributions. *J. Multivariate Anal.*, 116:109–129.

Gaenssler, P. and Stute, W. (1987). *Seminar on empirical processes*, volume 9 of *DMV Seminar*. Birkhäuser Verlag, Basel.

Genest, C. and Ghoudi, K. (1994). Une famille de lois bidimensionnelles insolite. *C. R. Acad. Sci. Paris Sér. I Math.*, 318(4):351–354.

Genest, C., Mesfioui, M., and Nešlehová, J. G. (in press, 2019). On the asymptotic covariance of the multivariate empirical copula process. *Depend. Model.*, 7.

Genest, C., Nešlehová, J., and Ziegel, J. (2011). Inference in multivariate Archimedean copula models. *TEST*, 20:223–256.

Genest, C. and Nešlehová, J. (2007). A primer on copulas for count data. *Astin Bull.*, 37(2):475–515.

Genest, C., Nešlehová, J. G., and Rémillard, B. (2017). Asymptotic behavior of the empirical multilinear copula process under broad conditions. *J. Multivariate Anal.*, 159:82–110.

Genest, C. and Rivest, L.-P. (1989). A characterization of gumbel's family of extreme value distributions. *Statistics & Probability Letters*, 8(3):207–211.

Genest, C. and Rivest, L.-P. (1993). Statistical inference procedures for bivariate Archimedean copulas. *J. Amer. Statist. Assoc.*, 88:1034–1043.

Genest, C. and Segers, J. (2009). Rank-based inference for bivariate extreme-value copulas. *Ann. Statist.*, 37:2990–3022.

Genest, C. and Segers, J. (2010). On the covariance of the asymptotic empirical copula process. *J. Multivariate Anal.*, 101(8):1837–1845.

Ghoudi, K., Khoudraji, A., and Rivest, L.-P. (1998). Propriétés statistiques des copules de valeurs extrêmes bidimensionnelles. *Canad. J. Statist.*, 26:187–197.

Gnedenko, B. (1943). Sur la distribution limite du terme maximum d'une série aléatoire. *Ann. of Math. (2)*, 44:423–453.

Gudendorf, G. and Segers, J. (2011). Nonparametric estimation of an extreme-value copula in arbitrary dimensions. *J. Multivariate Anal.*, 102:37–47.

Gudendorf, G. and Segers, J. (2012). Nonparametric estimation of multivariate extreme-value copulas. *J. Statist. Plann. Inference*, 142:3073–3085.

Hall, P. and Tajvidi, N. (2000). Distribution and dependence-function estimation for bivariate extreme-value distributions. *Bernoulli*, 6:835–844.

Hofert, M. (2008). Sampling Archimedean copulas. *Comput. Statist. Data Anal.*, 52(12):5163–5174.

Hofert, M., Huser, R., and Prasad, A. (2018). Hierarchical Archimax copulas. *J. Multivariate Anal.*, 167:195–211.

Hofert, M. and Maechler, M. (2016). Parallel and other simulations in R made easy: An end-to-end study. *J. Statist. Software*, 69:1–44.

Hoffman-Jorgensen, J. (1994). *Probability With a View Towards Statistics*, volume I. CRC Press.

Hsing, T. (1989). Extreme value theory for multivariate stationary sequences. *J. Multivariate Anal.*, 29(2):274–291.

Huang, X. (1992). *Statistics of bivariate extreme values*. Tinbergen Institute Research Series 22. Thesis (Ph.D.)–Erasmus University Rotterdam.

Huser, R., Opitz, T., and Thibaud, E. (2017). Bridging asymptotic independence and dependence in spatial extremes using Gaussian scale mixtures. *Spat. Stat.*, 21:166–186.

Hüsler, J. (1990). Multivariate extreme values in stationary random sequences. *Stochastic Process. Appl.*, 35(1):99–108.

Jenkinson, A. F. (1955). The frequency distribution of the annual maximum (or minimum) values of meteorological elements. *Quarterly Journal of the Royal Meteorological Society*, 81(348):158–171.

Joe, H. (2014). *Dependence Modeling With Copulas*. CRC Press, Boca Raton, FL.

Kallenberg, O. (2002). *Foundations of modern probability*. Probability and its Applications (New York). Springer-Verlag, New York, second edition.

Kaufman, L. and Rousseeuw, P. J. (1990). *Finding groups in data: an introduction to cluster analysis*. John Wiley & Sons.

Kendall, M. G. and Babington Smith, B. (1940). On the method of paired comparisons. *Biometrika*, 31:324–345.

Kimberling, C. H. (1974). A probabilistic interpretation of complete monotonicity. *Aequationes Math.*, 10:152–164.

Kojadinovic, I., Segers, J., and Yan, J. (2011). Large-sample tests of extreme-value dependence for multivariate copulas. *Canad. J. Statist.*, 39:703–720.

Larsson, M. and Nešlehová, J. (2011). Extremal behavior of Archimedean copulas. *Adv. in Appl. Probab.*, 43:195–216.

Leadbetter, M. R., Lindgren, G., and Rootzén, H. (1983). *Extremes and related properties of random sequences and processes*. Springer Series in Statistics. Springer-Verlag, New York.

Ledford, A. W. and Tawn, J. A. (1996). Statistics for near independence in multivariate extreme values. *Biometrika*, 83:169–187.

Mai, J. and Scherer, M. (2014). *Financial engineering with copulas explained*. Springer.

Malov, S. (2001). On finite-dimensional Archimedean copulas. In Balakrishnan, N., Ibragimov, I., and Nevzorov, V., editors, *Asymptotic Methods in Probability and Statistics with Applications*, pages 19–35, Basel. Birkhäuser.

McNeil, A. J. (2008). Sampling nested Archimedean copulas. *J. Stat. Comput. Simul.*, 78(5-6):567–581.

McNeil, A. J., Frey, R., Embrechts, P., et al. (2005). *Quantitative risk management: Concepts, techniques and tools*, volume 3. Princeton university press Princeton.

McNeil, A. J. and Nešlehová, J. (2009). Multivariate Archimedean copulas, d-monotone functions and ℓ_1-norm symmetric distributions. *Ann. Statist.*, 37:3059–3097.

McNeil, A. J. and Nešlehová, J. (2010). From Archimedean to Liouville copulas. *J. Multivariate Anal.*, 101(8):1772–1790.

Mesiar, R. and Jágr, V. (2013). d-dimensional dependence functions and Archimax copulas. *Fuzzy Sets and Systems*, 228:78–87.

Mises, R. v. (1936). La distribution de la plus grande de n valeurs. *Rev. Math. Union Interbalcanique*, 1:141–160.

Morillas, P. M. (2005). A characterization of absolutely monotonic (Δ) functions of a fixed order. *Publ. Inst. Math. (Beograd) (N.S.)*, 78(92):93–105.

Murphy, O. (2018). *Extremal modeling of dependent and non-stationary series and joint modeling of extreme rainfall at multiple sites*. McGill University Libraries. Thesis (Ph.D.)–McGill University.

Nelsen, R. B. (2006). *An Introduction to Copulas*. Springer, New York, second edition.

Pickands, III, J. (1975). Statistical inference using extreme order statistics. *Ann. Statist.*, 3:119–131.

Pickands, III, J. (1981). Multivariate extreme value distributions. In *Proceedings of the 43rd session of the International Statistical Institute, Vol. 2 (Buenos Aires, 1981)*, volume 49, pages 859–878, 894–902.

Resnick, S. I. (1987). *Extreme Values, Regular Variation, and Point Processes*. Springer-Verlag, New York.

Ressel, P. (2013). Homogeneous distributions—and a spectral representation of classical mean values and stable tail dependence functions. *J. Multivariate Anal.*, 117:246–256.

Rousseeuw, P. J. (1987). Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. *Journal of computational and applied mathematics*, 20:53–65.

Rüschendorf, L. (1976). Asymptotic distributions of multivariate rank order statistics. *Ann. Statist.*, 4(5):912–923.

Salvadori, G., De Michele, C., Kottegoda, N. T., and Rosso, R. (2007). *Extremes in nature: an approach using copulas*, volume 56. Springer Science & Business Media.

Segers, J. (2012). Asymptotics of empirical copula processes under non-restrictive smoothness assumptions. *Bernoulli*, 18:764–782.

Sklar, A. (1959). Fonctions de répartition à n dimensions et leurs marges. *Publ. Inst. Statist. Univ. Paris*, 8:229–231.

Stute, W. (1984). The oscillation behavior of empirical processes: the multivariate case. *Ann. Probab.*, 12(2):361–379.

Tsukahara, H. (2000). Empirical copulas and some applications. Technical Report 27, Seijo University: Tokyo, Japan.

van der Vaart, A. W. (1998). *Asymptotic statistics*, volume 3 of *Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge University Press, Cambridge.

van der Vaart, A. W. and Wellner, J. A. (1996). *Weak Convergence and Empirical Processes, With Applications to Statistics*. Springer, New York.

Vatan, P. (1985). Max-infinite divisibility and max-stability in infinite dimensions. In *Probability in Banach spaces, V (Medford, Mass., 1984)*, volume 1153 of *Lecture Notes in Math.*, pages 400–425. Springer, Berlin.

Vettori, S., Huser, R., Segers, J., and Genton, M. G. (2019). Bayesian model averaging over tree-based dependence structures for multivariate extremes. *Journal of Computational and Graphical Statistics*, (just-accepted):1–37.

Wadsworth, J. L., Tawn, J. A., Davison, A. C., and Elton, D. M. (2017). Modelling across extremal dependence classes. *J. R. Stat. Soc. Ser. B. Stat. Methodol.*, 79:149–175.

Williamson, R. E. (1956). Multiply monotone functions and their Laplace transforms. *Duke Math. J.*, 23:189–207.

Zhang, D., Wells, M. T., and Peng, L. (2008). Nonparametric estimation of the dependence function for a multivariate extreme value distribution. *J. Multivariate Anal.*, 99:577–588.

Modélisation de la dépendance entre pré-extrêmes

Résumé: Cette thèse développe des techniques d'inférence pour copule Archimax $C_{\psi,\ell}$. Des conditions sont dérivées afin que ψ et ℓ soient identifiables, de sorte qu'une approche d'inférence semi-paramétrique puisse être développée. Deux estimateurs non paramétriques de ℓ et un estimateur de ψ basé sur les moments, supposant que ce dernier appartient à une famille paramétrique, sont avancés. Le comportement asymptotique de ces estimateurs est ensuite établi sous des hypothèses de régularité et la performance en échantillon fini est évaluée par le biais d'une étude de simulation. Une construction hiérarchique qui généralise les copules Archimax est proposée afin d'apporter davantage de flexibilité. Le comportement extrême de ce nouveau modèle de dépendance est ensuite étudié. La copule Archimax est ensuite utilisée pour analyser des maxima mensuels de précipitations. L'estimateur non paramétrique de ℓ révèle une dépendance extrême asymétrique entre les stations, ce qui reflète le déplacement des orages dans la région. Une application du modèle Archimax hiérarchique à un jeu de données de précipitations contenant 155 stations est ensuite présentée, dans laquelle des groupes de stations asymptotiquement dépendantes sont déterminés via un algorithme de "clustering" spécifiquement adapté au modèle.

Mots clés: modélisation de la dépendance; extrêmes; pré-extrêmes; copules; inférence semi paramétrique; modélisation hiérarchique; processus empiriques; asymptotique; convergence faible; précipitation extrême.

Dependence modeling for pre-asymptotic extremes

Abstract: This thesis develops inference techniques for Archimax copulas, which are denoted $C_{\psi,\ell}$. Conditions under which ψ and ℓ are identifiable are derived so that a semiparametric approach for inference can be developed. Two nonparametric estimators of ℓ and a moment-based estimator of ψ, which assumes that the latter belongs to a parametric family, are proposed. The asymptotic behavior of the estimators is then established under broad regularity conditions; performance in small samples is assessed through a comprehensive simulation study. Archimax copulas are then generalized to a clustered constructions in order to bring in more flexibility. The extremal behavior of this new dependence model is derived. Finally, the methodology proposed herein is illustrated on precipitation data. First, a trivariate Archimax copula is used to analyze monthly rainfall maxima. The nonparametric estimator of ℓ reveals asymmetric extremal dependence between the stations, which reflects heavy precipitation patterns in the area. An application of the clustered Archimax model to a precipitation dataset containing 155 stations is then presented, where groups of asymptotically dependent stations are determined via a specifically tailored clustering algorithm.

Keywords: dependence modeling; extremes; pre-extremes; copulas; semiparametric inference; hierarchical modeling; empirical processes; asymptotics; weak convergence; extreme precipitation.

Image en couverture : Papagaio-ajuretê. Dayana Moraes.