Working Paper in Economics

202503

Copula-based Measures of Extreme Dependence

Mohamed Doukali[1]
Xiaojun Song[2]
Abderrahim Taamouti[3]

[1] University of Liverpool Management School, m.doukali@liverpool.ac.uk
[2] Peking University, sxj@gsm.pku.edu.cn
[3] University of Liverpool Management School, abderrahim.taamouti@liverpool.ac.uk

https://www.liverpool.ac.uk/economics/

Copula-based Measures of Extreme Dependence

Mohamed Doukali † [1], Xiaojun Song ‡[2] and Abderrahim Taamouti §[1]

[1]University of Liverpool Management School.

[2]Guanghua School of Management

Version: 3 April 2025

Abstract

We introduce measures of extreme dependence using Kullback-Leibler relative entropy, which we define in terms of copula densities. To estimate these measures, we employ the Bernstein copula density estimator, providing consistent nonparametric estimators and deriving their Bahadur-type representations. Subsequently, we establish the asymptotic distribution of a test for symmetric dependence, constructed using these measures. We explore the properties of this test under both global and local alternatives. Additionally, we demonstrate the validity of a bootstrap-based test that can be employed in finite-sample settings to assess symmetric dependence. A Monte Carlo simulation study shows that both the asymptotic and bootstrap-based tests exhibit good finite-sample size and power properties across various data-generating processes and sample sizes. Finally, we present an empirical application that highlights the practical utility of the extreme dependence measures. Specifically, we quantify the degree of extreme dependence between the US financial market and several developed and emerging financial markets.

Keywords: Measuring extreme dependence, asymmetric dependence, Bernstein copula density, international stock market co-movements.

Journal of Economic Literature classification: C12, C14, G15.

†Department of Economics, University of Liverpool Management School. E-mail: M.Doukali@liverpool.ac.uk.

‡Department of Business Statistics and Econometrics, Guanghua School of Management and Center for Statistical Science, Peking University, Beijing, 100871, China. E-mail: sxj@gsm.pku.edu.cn. Financial support from the National Natural Science Foundation of China (Grant No. 71973005) is acknowledged.

§Department of Economics, University of Liverpool Management School. E-mail: abderrahim.taamouti@liverpool.ac.uk.

1 Introduction

Numerous empirical studies have established that financial markets exhibit asymmetric dependence, meaning they tend to be more interdependent during market downturns and less so during market upturns. Accounting for this type of dependence, and its degree, can significantly enhance portfolio allocation and risk management strategies; see Ang and Chen (2002), among others. The critical importance of understanding the nature of dependence—whether symmetric or asymmetric—between asset returns has encouraged the development of various parametric and nonparametric tests to assess symmetric dependence in financial markets; see Longin and Solnik (2001); Ang and Bekaert (2002); Ang and Chen (2002); Patton (2004); Hong et al. (2007); Jiang, Wu and Zhou (2018); and Song and Taamouti (2017). However, these tests fall short of quantifying the degree of dependence, as they primarily indicate the presence or absence of dependence. Consequently, a strong asymmetric dependence might not be statistically significant at a given significance level, while a statistically significant asymmetric dependence might not be "strong" from an economic perspective or relevant for investment decision-making.

In this paper, we propose nonparametric measures for quantifying extreme dependence. These measures are capable of capturing both linear and nonlinear forms of extreme dependence. For instance, they can assess the degree of co-movement between the U.S. market and other international markets, thereby identifying which markets are more or less vulnerable to potential economic or financial crises originating in the U.S. Additionally, as will be demonstrated later, these measures can also be employed to develop tests for symmetric dependence.

In a recent study, Jiang, Maasoumi, Pan and Wu (2018) interested in measuring general asymmetric dependence between two random variables, building upon the exceedance correlation measure introduced by Hong, Tu and Zhou (2007). Their approach was applied to test the hypothesis of symmetric dependence. Specifically, they proposed a two-step procedure to test for general asymmetric exceedance dependence in stock returns. However, beyond the fact that their approach is based on the probability density function, their study lacks theoretical results concerning the consistency of the estimators for their measures and does not derive the asymptotic distributions. In this paper, we address these gaps by presenting measures of extreme dependence based on copula density functions and rigorously examine the consistency and establish the asymptotic distributions of their nonparametric estimators.

Formally, we consider two univariate random variables, X and Y, with means $E(X) = \mu_X$ and $E(Y) = \mu_Y$, and variances $Var(X) = \sigma_X^2$ and $Var(Y) = \sigma_Y^2$, respectively. In the following, we assume that X and Y have been standardized to have a mean of zero and a variance of one. Using Kullback-Leibler relative entropy (Kullback and Leibler, 1951), Jiang, Maasoumi, Pan and Wu (2018) define the following measures of lower and upper tail exceedance dependence for a given positive exceedance level c:

$$\rho^{-}(c)=\int_{-\infty}^{-c}\int_{-\infty}^{-c} f_{XY}(x,y)\log\left[\frac{f_{XY}(x,y)}{f_X(x)\,f_Y(y)}\right]dxdy,$$

$$\rho^{+}(c)=\int_{c}^{+\infty}\int_{c}^{+\infty} f_{XY}(x,y)\log\left[\frac{f_{XY}(x,y)}{f_X(x)\,f_Y(y)}\right]dxdy,$$

where, $f_{XY}(x,y)$ represents the joint probability density function of $(X,Y)'$, while $f_X(x)$ and $f_Y(y)$ denote the marginal probability density functions of X and Y, respectively. To estimate $\rho^{-}(c)$ and $\rho^{+}(c)$, Jiang, Maasoumi, Pan and Wu (2018) propose replacing $f_{XY}(x,y)$, $f_X(x)$, and $f_Y(y)$ with their nonparametric estimates obtained through kernel methods. However, they did not explore the theoretical properties, such as consistency and asymptotic distribution, of these estimators. Instead, they relied on Monte Carlo simulations to evaluate the performance of their test based on Kullback-Leibler relative entropy.

In this paper, we first propose measures of extreme dependence using Bernstein copula density functions, as opposed to the probability density functions employed in Jiang, Maasoumi, Pan and Wu (2018). Bernstein copula density functions are particularly advantageous due to their ability to address boundary bias issues, a common problem in kernel methods like those used by Jiang, Maasoumi, Pan and Wu (2018). Subsequently, we conduct a rigorous analysis of the asymptotic properties of the nonparametric estimators for our proposed measures.

Specifically, we begin by expressing these measures in terms of the copula function. In other words, using Sklar (1959)'s theorem, the measures can be reformulated as follows:

$$\rho^{-}(c)=\int_{-\infty}^{-c}\int_{-\infty}^{-c} f_{XY}(x,y)\log\left[c\left(F_X(x),F_Y(y)\right)\right]dxdy, \tag{1.1}$$

and

$$\rho^{+}(c)=\int_{c}^{+\infty}\int_{c}^{+\infty} f_{XY}(x,y)\log\left[c\left(F_X(x),F_Y(y)\right)\right]dxdy, \tag{1.2}$$

where $c(\cdot,\cdot)$ is the copula density of $(X,Y)'$ and $F_X(X)$ and $F_Y(Y)$ are the cumulative distribution functions of X and Y, respectively. We can further re-write (1.1) and (1.2) as follows:

$$\rho^{-}(c) = E\left\{\log\left[c\left(F_X(X), F_Y(Y)\right)\right] \mathscr{I}(X \le -c, Y \le -c)\right\},$$
$$\rho^{+}(c) = E\left\{\log\left[c\left(F_X(X), F_Y(Y)\right)\right] \mathscr{I}(X \ge c, Y \ge c)\right\}, \tag{1.3}$$

where $\mathscr{I}(X \le -c, Y \le -c)$ and $\mathscr{I}(X \ge c, Y \ge c)$ are indicator functions that we define in the next section [see Equation (2.2)].

Moreover, we present nonparametric estimators for the measures $\rho^{-}(c)$ and $\rho^{+}(c)$ as defined in (1.3), using the Bernstein copula density estimator introduced by Sancetta and Satchell (2004) and Bouezmarni, Rombouts and Taamouti (2010). We establish the consistency of these estimators and derive their Bahadur-type representations. Given that these measures can be employed to test for symmetric dependence, another significant contribution of this paper is the derivation of the asymptotic distributions for tests of symmetric dependence based on our proposed measures. Subsequently, we analyze the properties of these tests under both global and local alternatives.

Additionally, we propose a bootstrap-based test that can be applied in finite-sample contexts to assess the hypothesis of symmetric dependence. A Monte Carlo simulation study demonstrates that both the asymptotic and bootstrap-based tests exhibit robust finite-sample size and power properties across a range of data-generating processes and sample sizes. Finally, we provide an empirical application to demonstrate the practical utility of our extreme dependence measures, specifically by quantifying the degree of extreme dependence between the US financial market and several developed and emerging financial markets.

The remainder of this paper is structured as follows. Section 2 introduces the general framework for estimating lower and upper tail exceedance dependence measures using the Bernstein copula density. In Section 3, we explore the asymptotic properties of the estimators for tail dependence measures, specifically proving their consistency and deriving their Bahadur representations. Section 4 demonstrates how these measures and their estimators can be utilized to construct a test for symmetric dependence. We derive the asymptotic distribution of this test and analyze its power under both global and local alternatives. In Section 5, we propose an alternative bootstrap-based test for symmetric dependence and conduct a Monte Carlo simulation study to assess the finite-sample properties of both bootstrap and asymptotic tests. Section 6 is dedicated to an empirical application, with conclusions drawn from the results presented in Section 7. Finally, the tables of the empirical results and proofs are provided in Appendix A and Appendix B, respectively.

2 Framework

Consider two variables of interest, X and Y, and let $(Y_t, X_t) \in \mathbb{R}^2$, $t \in \mathbb{Z}$ be a strictly stationary stochastic process in $\mathbb{R}^2$. Denote by $f(x,y)$ the joint probability density function of X_t and Y_t, and by $c(F(X), F(Y))$ the corresponding copula density function, where $F(X)$ and $F(Y)$ are the cumulative distribution functions of X and Y, respectively. As discussed in the introduction, we are interested in providing measures of extreme dependence between X and Y using copula density. Thus, for a non-negative constant c, and following the work of Jiang, Maasoumi, Pan and Wu (2018), we use the copula function $c(F(X), F(Y))$ to define the following measures of lower and upper tail exceedance dependence:

$$\begin{aligned}
\rho^-(c) &= E\left\{\log\left[c\left(F_X(X), F_Y(Y)\right)\right] \mathscr{I}(X \leq -c, Y \leq -c)\right\}, \\
\rho^+(c) &= E\left\{\log\left[c\left(F_X(X), F_Y(Y)\right)\right] \mathscr{I}(X \geq c, Y \geq c)\right\},
\end{aligned} \tag{2.1}$$

where $\mathscr{I}(X \leq -c, Y \leq -c)$ and $\mathscr{I}(X \geq c, Y \geq c)$ are indicator functions that we define as follows:

$$\begin{aligned}
\mathscr{I}(X \leq -c, Y \leq -c) &= \begin{cases} 1, \text{ if } X \leq -c \text{ and } Y \leq -c \\ 0, \text{ otherwise,} \end{cases} \\
\mathscr{I}(X \geq c, Y \geq c) &= \begin{cases} 1, \text{ if } X \geq c \text{ and } Y \geq c \\ 0, \text{ otherwise.} \end{cases}
\end{aligned} \tag{2.2}$$

In this section, we discuss the estimation of the above measures. Considering the expressions of the measures in (2.1), it appears that natural estimators of $\rho^-(c)$ and $\rho^+(c)$ can be obtained by replacing the unknown copula $c(F_X(X), F_Y(Y))$ with its nonparametric estimate from a finite sample, and the theoretical expectation "E" with its empirical analogue $T^{-1}\sum_{t=1}^T$. In what follows, we use the Bernstein copula density as an estimator of $c(F_X(X), F_Y(Y))$ due to its well-known attractive properties, such as addressing the boundary bias problem that affects standard kernel methods like the one used in Jiang, Maasoumi, Pan and Wu (2018). Formally, we consider the following nonparametric estimators of $\rho^-(c)$ and $\rho^+(c)$:

$$\begin{aligned}
\hat{\rho}^-(c) &= \tfrac{1}{T}\textstyle\sum_{t=1}^T \log\left[c_{k,T}(\mathbf{V_t})\right] \mathscr{I}(X_t \leq -c, Y_t \leq -c), \\
\hat{\rho}^+(c) &= \tfrac{1}{T}\textstyle\sum_{t=1}^T \log\left[c_{k,T}(\mathbf{V_t})\right] \mathscr{I}(X_t \geq c, Y_t \geq c),
\end{aligned} \tag{2.3}$$

where $c_{k,T}(\cdot)$ is the Bernstein copula density estimator of $c\left(F_X(X), F_Y(Y)\right)$ that we define below.

We now define the Bernstein copula density estimator, which we use to construct our nonparametric estimators for the measures of tail dependence between X and Y. Without loss of generality, we focus on the case where X and Y are both univariate variables. The more general multidimensional case can be analyzed similarly, albeit with slightly more complex notation. In what follows, we denote $\mathbf{u} = (u_1, u_2)' = \left(F_X(X), F_Y(Y)\right)' \in [0,1]^2$. Thus, the copula density $c(\mathbf{u})$ is defined as follows:

$$c(\mathbf{u}) = \partial^2 C(\mathbf{u}) / \partial u_1 \partial u_2,$$

where $C(\cdot)$ is the copula function. The Bernstein copula function is absolutely continuous, hence the Bernstein copula density is defined as:

$$c_k(\mathbf{u}) = \sum_{v_1=0}^{k} \sum_{v_2=0}^{k} C\left(\frac{v_1}{k}, \frac{v_2}{k}\right) \prod_{j=1}^{2} P'_{v_j,k}(u_j),$$

where $P'_{v_j,k}(u_j)$ is the derivative of $P_{v_j,k}(u_j)$ with respect to u_j, with $P_{v_j,k}(u_j)$ the binomial probability mass with parameters v_j and k evaluated at u_j. The Bernstein estimator of the copula density is given by

$$c_{k,T}(\mathbf{u}) = \sum_{v_1=0}^{k} \sum_{v_2=0}^{k} C_T\left(\frac{v_1}{k}, \frac{v_2}{k}\right) \prod_{j=1}^{2} P'_{v_j,k}(u_j),$$

where $C_T(\cdot)$ is the empirical copula and T is the number of observations, see e.g. Sancetta and Satchell (2004) and Bouezmarni, Rombouts and Taamouti (2010). The Bernstein copula density estimator can be rewritten as follows [see Bouezmarni, Rombouts and Taamouti (2010)]:

$$c_{k,T}(\mathbf{u}) = \frac{1}{T} \sum_{t=1}^{T} K_k(\mathbf{u}, \mathbf{V}_t), \text{ for } \mathbf{u} \in [0,1]^2, \tag{2.4}$$

with

$$K_k(\mathbf{u}, \mathbf{V}_t) = k^2 \sum_{v_1=0}^{k-1} \sum_{v_2=0}^{k-1} \mathscr{I}\left(\mathbf{V}_t \in A_k(\mathbf{v})\right) \prod_{j=1}^{2} P_{v_j,k-1}(u_j),$$

where $P_{v_j,k-1}(u_j)$ is the binomial probability mass with parameters v_j and $k-1$ evaluated at u_j, and

$$\mathbf{V}_t = \left(F_{XT}(X_t), F_{YT}(Y_t)\right)',$$

with $F_{XT}(x) = T^{-1} \sum_{t=1}^{T} \mathscr{I}(X_t \leq x)$ and $F_{YT}(y) = T^{-1} \sum_{t=1}^{T} \mathscr{I}(Y_t \leq y)$, respectively, the empirical distribution functions of X_t and Y_t for $t = 1, \ldots, T$, and $A_k(\mathbf{v}) = \left[\frac{v_1}{k}, \frac{v_1+1}{k}\right] \times \left[\frac{v_2}{k}, \frac{v_2+1}{k}\right]$ for $\mathbf{v} = (v_1, v_2)'$, with the integer k that plays the role of bandwidth parameter. The Bernstein copula density estimator was first proposed and investigated by Sancetta and Satchell (2004) for independent and identically distributed (i.i.d.) data. Later, Bouezmarni, Rombouts and Taamouti (2010) applied Bernstein polynomials to estimate the copula density for time series data with α-mixing dependence. They provided the asymptotic properties of the Bernstein copula density estimator for α-mixing data, including asymptotic bias and variance, uniform almost sure (a.s.) convergence, and asymptotic normality. More recently, Janssen, Swanepoel and Veraverbeke (2014) revisited this estimator, establishing its asymptotic normality under i.i.d. conditions. In the following, we study the asymptotic properties (consistency and asymptotic distribution) of the nonparametric estimators $\hat{\rho}^{-}(c)$ and $\hat{\rho}^{+}(c)$.

3 Asymptotic properties of $\hat{\rho}^{-}(c)$ and $\hat{\rho}^{+}(c)$

This section aims to investigate the asymptotic properties of the nonparametric estimators for the lower and upper tail exceedance dependence measures, denoted as $\hat{\rho}^{-}(c)$ and $\hat{\rho}^{+}(c)$, respectively. These estimators are crucial for understanding the extremal dependence structure in tail regions.

Before proceeding with the asymptotic analysis, we introduce a set of standard assumptions, which will serve as the foundation for our theoretical results. These assumptions ensure the regularity conditions necessary for deriving consistency properties of the estimators of our tail exceedance dependence measures as well as establishing their Bahadur representations, which can be used to obtain the asymptotic normality of the estimators.

Assumption A1: The process $(Y_t, X_t) \in \mathbb{R}^2$, $t \in \mathbb{Z}$ is a strictly stationary β-mixing process with coefficient $\beta_l = O(\rho^l)$, for some $0 < \rho < 1$.

Assumption A2: We assume that $k \to \infty$ together with $T^{-1/2}k^{3/2}\log\log^2(T) \to 0$ when $T \mapsto \infty$.

These assumptions are widely used in the literature, including in Bouezmarni, Rombouts and Taamouti (2012) and Belalia, Bouezmarni, Lemyre and Taamouti (2017), where a detailed discussion can also be found. Based on the results of Belalia, Bouezmarni, Lemyre and Taamouti (2017), the following proposition establishes the consistency of the measures $\hat{\rho}^{-}(c)$ and $\hat{\rho}^{+}(c)$. The proof of Proposition 1 follows from the proof of Theorem 2 in Belalia et al. (2017).

Proposition 1. *Under Assumptions **A1** and **A2**, the estimators $\hat{\rho}^{-}(c)$ and $\hat{\rho}^{+}(c)$ in (2.3) converges in probability to the true measures of extreme dependence $\rho^{-}(c)$ and $\rho^{+}(c)$, respectively.*

We now present the Bahadur representations for the estimators $\hat{\rho}^{-}(c)$ and $\hat{\rho}^{+}(c)$. Before doing so, we define

$$C_{u_l}(u_1,u_2) := \partial C(u_1,u_2)/\partial u_l, \text{ for } l=1,2,$$

as the first-order partial derivatives of the copula distribution function $C(u_1,u_2)$ over $(u_1,u_2)\in(0,1)^2$. The following theorem establishes the Bahadur representations of $\hat{\rho}^-(c)$ and $\hat{\rho}^+(c)$ [see the proof of Theorem 1 in Appendix B].

Theorem 1. *: Suppose Assumptions **A1** and **A2** hold. Then, for any given positive constant c, as $T\to\infty$, we have*

$$\begin{aligned}\sqrt{T}\left[\hat{\rho}^-(c)-\rho^-(c)\right] &= \tfrac{1}{\sqrt{T}}\textstyle\sum_{t=1}^{T}\left[\varepsilon_{1t}^-(c)+\varepsilon_{2t}^-(c)\right]+o_p(1),\\ \sqrt{T}\left[\hat{\rho}^+(c)-\rho^+(c)\right] &= \tfrac{1}{\sqrt{T}}\textstyle\sum_{t=1}^{T}\left[\varepsilon_{1t}^+(c)+\varepsilon_{2t}^+(c)\right]+o_p(1)\end{aligned}$$

where

$$\varepsilon_{1t}^-(c)=\log\left[c\left(F_X(X_t),F_Y(Y_t)\right)\right]\mathscr{I}(X_t\le -c,Y_t\le -c)-\rho^-(c),$$

$$\begin{aligned}\varepsilon_{2t}^-(c)=\;&\mathscr{I}(X_t\le -c,Y_t\le -c)-E\left[\mathscr{I}(X_t\le -c,Y_t\le -c)\right]\\ &-C_{u_1}\left(F_X(-c),F_Y(-c)\right)\left[\mathscr{I}(X_t\le -c)-F_X(-c)\right]\\ &-C_{u_2}\left(F_X(-c),F_Y(-c)\right)\left[\mathscr{I}(Y_t\le -c)-F_Y(-c)\right],\end{aligned}$$

$$\varepsilon_{1t}^+(c)=\log\left[c\left(F_X(X_t),F_Y(Y_t)\right)\right]\mathscr{I}(X_t\ge c,Y_t\ge c)-\rho^+(c),$$

$$\begin{aligned}\varepsilon_{2t}^+(c)=\;&\mathscr{I}(X_t\ge c,Y_t\ge c)-E\left[\mathscr{I}(X_t\ge c,Y_t\ge c)\right]\\ &+C_{u_1}\left(1-F_X(c),1-F_Y(c)\right)\left[\mathscr{I}(X_t\le c)-F_X(c)\right]\\ &+C_{u_2}\left(1-F_X(c),1-F_Y(c)\right)\left[\mathscr{I}(Y_t\le c)-F_Y(c)\right].\end{aligned}$$

In the Bahadur representations of $\hat{\rho}^-(c)$ and $\hat{\rho}^+(c)$, although the expressions may seem complex, their interpretation is relatively straightforward. The first terms, $\varepsilon_{1t}^-(c)$ and $\varepsilon_{1t}^+(c)$, represent the standard Bahadur terms when the true copula density function $c(u_1,u_2)$ is known. In contrast, the second terms, $\varepsilon_{2t}^-(c)$ and $\varepsilon_{2t}^+(c)$, reflect the non-negligible estimation effects that arise from replacing the unknown true copula density $c(u_1,u_2)$ with the Bernstein copula density approximation $c_{k,T}(u_1,u_2)$.

Moreover, these estimation effects can be further traced to distinct sources. Specifically, the initial terms in $\varepsilon_{2t}^-(c)$ and $\varepsilon_{2t}^+(c)$ result from the need to replace the unknown copula distribution function $C(u_1,u_2)$ with the empirical Bernstein copula distribution estimator. The remaining terms stem from the fact that the marginal distributions $F_X(x)$ and $F_Y(y)$ are also unknown and must be substituted with their empirical distribution functions (EDFs), $F_{XT}(x)$ and $F_{YT}(y)$.

Interestingly, by applying the Central Limit Theorem, the results in Theorem 1 demonstrate that, despite the fact that our estimators are model-free and entirely nonparametric, we still achieve the standard parametric rate of convergence.

4 Test of asymmetric dependence

The measures of tail dependence described previously can be used to construct tests for symmetric dependence at the extremes. Specifically, if there exists symmetric dependence between the random variables X and Y at a given exceedance level $c > 0$, then the measures of lower and upper tail exceedance dependence, $\rho^-(c)$ and $\rho^+(c)$, will be equal. This equality forms the basis for testing the null hypothesis of symmetric tail dependence. To formalize, for a pre-specified exceedance level c, we aim to test the null hypothesis:

$$H_0(c) : \rho^-(c) = \rho^+(c), \tag{4.1}$$

against the alternative hypothesis:

$$H_1(c) : \rho^-(c) \neq \rho^+(c). \tag{4.2}$$

Under this null hypothesis, if the two measures are indeed equal, it suggests that the dependence between X and Y is symmetric at the threshold c. To perform this test, we use the empirical estimators $\rho^-(c)$ and $\rho^+(c)$ and assess whether their observed difference is statistically significant.

In the following analysis, we employ a two-sided t-type test to assess the null hypothesis $H_0(c)$ against the alternative $H_1(c)$. This test is based on the distance between the measures of tail dependence $\rho^-(c)$ and $\rho^+(c)$. To determine the asymptotic distribution of our test statistic, we first investigate the asymptotic behavior of the difference $\hat{\rho}^-(c) - \hat{\rho}^+(c)$. Theorem 2 provides an asymptotic decomposition of the term $\sqrt{T}\,[\hat{\rho}^-(c) - \hat{\rho}^+(c)]$, which allows us to derive its asymptotic distribution. For a detailed proof of Theorem 2, please refer to Appendix B.

Theorem 2. **:** *Let Assumptions **A1** and **A2** hold. Then, under the null hypothesis in* (4.1)*, for given*

$c > 0$, as $T \to \infty$, we have

$$\begin{aligned}\sqrt{T}\left[\hat{\rho}^{-}(c)-\hat{\rho}^{+}(c)\right] =& \frac{1}{\sqrt{T}}\sum_{t=1}^{T}\left\{\log\left[c\left(F_X\left(X_t\right),F_Y\left(Y_t\right)\right)\right]\left[\mathscr{I}\left(X_t \le -c, Y_t \le -c\right)-\mathscr{I}\left(X_t \ge c, Y_t \ge c\right)\right]\right. \\ & \left. +\left[\varepsilon_{2t}^{-}(c)-\varepsilon_{2t}^{+}(c)\right]\right\}+o_p(1) \\ :=& \frac{1}{\sqrt{T}}\sum_{t=1}^{T}u_t(c)+o_p(1). \end{aligned} \tag{4.3}$$

where,

$$u_t(c)=\left\{\log\left[c\left(F_X\left(X_t\right),F_Y\left(Y_t\right)\right)\right]\left[\mathscr{I}\left(X_t \le -c, Y_t \le -c\right)-\mathscr{I}\left(X_t \ge c, Y_t \ge c\right)\right]+\left[\varepsilon_{2t}^{-}(c)-\varepsilon_{2t}^{+}(c)\right]\right\}$$

Furthermore, under the null hypothesis,

$$\sqrt{T}\left[\hat{\rho}^{-}(c)-\hat{\rho}^{+}(c)\right] \to_d N\left(0,\sigma^2(c)\right),$$

where the asymptotic variance $\sigma^2(c)$ is given by

$$\sigma^2(c)=E\left(u_1^2(c)\right)+2\sum_{j=1}^{\infty}E\left[u_1(c)\,u_{1+j}(c)\right].$$

Notice that for any given exceedance level $c > 0$, if the long-run variance $\sigma^2(c)$ was known, one could directly use the quantity $\sqrt{T}\left[\hat{\rho}^{-}(c)-\hat{\rho}^{+}(c)\right]$ to construct a two-sided t-type test statistic, with the following limiting null distribution:

$$t(c) := \frac{\sqrt{T}\left[\hat{\rho}^{-}(c)-\hat{\rho}^{+}(c)\right]}{\sigma(c)} \to_d N(0,1),$$

whose asymptotic critical values are available in most econometric and statistical textbooks. However, in practice, the asymptotic variance $\sigma^2(c)$ is unknown, rendering the test statistic $t(c)$ infeasible. Thus, a proper estimator of $\sigma^2(c)$ is required. For a given weakly consistent estimator, $\hat{\sigma}^2(c) \to_p \sigma^2(c)$ [see e.g. Liu and Wu (2010)], our proposed feasible t-type test statistic is defined as

$$\hat{t}(c) := \frac{\sqrt{T}\left[\hat{\rho}^{-}(c)-\hat{\rho}^{+}(c)\right]}{\hat{\sigma}(c)}, \tag{4.4}$$

whose asymptotic distribution, as $T \to \infty$, remains unchanged under the null hypothesis and follows a standard normal distribution. Notice that when the process $\{u_t(c)\}$, defined in (4.3), is a martingale difference, the variance $\sigma^2(c)$ reduces to $E\left(u_1^2(c)\right)$ and can be consistently estimated by

$$\frac{1}{T}\sum_{t=1}^{T}\hat{u}_t^2(c),$$

where $\hat{u}_t(c)$ is defined as follows:

$$\hat{u}_t(c) = \log\left[c_{k,T}\left(F_{XT}(X_t), F_{YT}(Y_t)\right)\right]\left[\mathscr{I}(X_t \leq -c, Y_t \leq -c) - \mathscr{I}(X_t \geq c, Y_t \geq c)\right] + \left[\hat{\varepsilon}_{2t}^{-}(c) - \hat{\varepsilon}_{2t}^{+}(c)\right],$$

with

$$\begin{aligned}\hat{\varepsilon}_{2t}^{-}(c) =& \mathscr{I}(X_t \leq -c, Y_t \leq -c) - \frac{1}{T}\sum_{t=1}^{T}\mathscr{I}(X_t \leq -c, Y_t \leq -c)\\ &- \widehat{C}_{u_1,T}\left(F_{XT}(-c), F_{YT}(-c)\right)\left[\mathscr{I}(X_t \leq -c) - F_{XT}(-c)\right]\\ &- \widehat{C}_{u_2,T}\left(F_{XT}(-c), F_{YT}(-c)\right)\left[\mathscr{I}(Y_t \leq -c) - F_{YT}(-c)\right],\end{aligned}$$

and

$$\begin{aligned}\hat{\varepsilon}_{2t}^{+}(c) =& \mathscr{I}(X_t \geq c, Y_t \geq c) - \frac{1}{T}\sum_{t=1}^{T}\mathscr{I}(X_t \geq c, Y_t \geq c)\\ &+ \widehat{C}_{u_1,T}\left(1 - F_{XT}(c), 1 - F_{YT}(c)\right)\left[\mathscr{I}(X_t \leq c) - F_{XT}(c)\right]\\ &+ \widehat{C}_{u_2,T}\left(1 - F_{XT}(c), 1 - F_{YT}(c)\right)\left[\mathscr{I}(Y_t \leq c) - F_{YT}(c)\right].\end{aligned}$$

We now examine the consistency and asymptotic power of the feasible $\hat{t}(c)$ test in (4.4) for detecting local departures from the null hypothesis (4.1). Specifically, we consider local alternatives that converge to the null hypothesis at an appropriate rate. To this end, we introduce a sequence of Pitman-type local alternatives of the following form:

$$H_{1T}(c) : \rho^{-}(c) = \rho^{+}(c) + \frac{\gamma(c)}{\sqrt{T}}, \tag{4.5}$$

where $\gamma(c)$ is a finite positive constant representing the magnitude of the deviation of $\rho^{-}(c)$ from $\rho^{+}(c)$. Asymptotically (as $T \to \infty$), these terms converge to equality. The following proposition demonstrates that our test is consistent and exhibits non-trivial asymptotic local power against the sequence of Pitman local alternatives defined in (4.5), which converges to the null at the rate $T^{-1/2}$ [see the proof of Proposition 2 in Appendix B].

Proposition 2. *Let Assumptions **A1** and **A2** hold. Then, under the alternative hypothesis $H_1(c)$, for* *given $c > 0$, as $T \to \infty$, the $\hat{t}(c)$ test in (4.4) consistent, i.e.*

$$Pr\left(\left|\hat{t}(c)\right| > b\right) \to 1$$

for $T \to \infty$ and any positive constant $b = o(T^{1/2})$. In addition, under the local alternatives $H_{1T}(c)$ in (4.5), we have:

$$\sqrt{T}\left[\hat{\rho}^{-}(c) - \hat{\rho}^{+}(c)\right] \to_d N\left(\gamma(c), \sigma^2(c)\right).$$

From Proposition 2, we can immediately conclude that the limiting distribution of our nonparametric test $\hat{t}(c)$ is nontrivially shifted whenever $\gamma(c) > 0$. Therefore, the proposed test is able to detect local alternatives that converge to the null hypothesis $H_0(c)$ at the parametric rate $T^{-1/2}$. The local power of the test increases with the magnitude of the deviation of $\gamma(c)$, and thus our test has nontrivial power against the local alternatives in (4.5), which approach the parametric rate $T^{-1/2}$ arbitrarily closely.

As shown in (4.4), the asymptotic null distribution of $\hat{t}(c)$ follows a standard normal distribution, with the corresponding critical values readily available. Additionally, the limiting distribution is independent of the nuisance parameter k used in estimating $\hat{\rho}^{-}(c)$ and $\hat{\rho}^{+}(c)$. However, it is well known that tests based on the asymptotic null distribution often suffer from size distortions and power losses in finite samples. The finite-sample performance of the asymptotic-based test is also expected to depend on the choice of k. To address these limitations, we propose implementing the test using a stationary bootstrap procedure.

A key advantage of the stationary bootstrap is that the resampled pseudo-time series remains stationary, due to the use of geometrically distributed random block sizes. The stationary bootstrap introduced by Politis and Romano (1994*b*) is a powerful block-resampling technique that has been widely applied in the time series literature; see also Politis and Romano (1994*a*) and Hwang and Shin (2011), among others. The following sections, which present Monte Carlo simulations and empirical applications, demonstrate that the stationary bootstrap procedure offers good approximations in finite samples and is robust to the choice of k.

5 Monte-Carlo simulations

In this section, we conduct a Monte Carlo simulation study to evaluate the performance of the bootstrap approach introduced at the end of the previous section. This approach offers a small-sample approximation of the asymptotic distribution as described in Theorem 2. Specifically, we evaluate its size and power in testing for symmetric dependence, as well as its effectiveness in reducing bias in the estimators of tail dependence measures.

5.1 Performance of Bootstrap-Based Test

Though the asymptotic-based test $\hat{t}(c)$ given by Theorem 2 is computationally efficient and straightforward to implement, the empirical size of the test statistic $\hat{t}(c)$ in small samples may differ significantly from the nominal significance level. Size distortion is almost inevitable in small samples, as confirmed by our unreported simulations. However, as emphasized at the end of the previous section, it is well known that certain bootstrap methods, such as the stationary block bootstrap (Kunsch, 1989), can effectively account for the dependence structure in weakly dependent time series data and help improve the performance of the asymptotic-based tests. Stationarity is maintained by randomly selecting the length of each block from a geometric distribution (Politis and Romano, 1994*b*), with the mean determined by the algorithm proposed in Politis and White (2004) and Patton et al. (2009).

In light of the above, this section proposes approximating the finite sample distribution of $\hat{t}(c)$ under the null hypothesis by using the distribution of

$$\hat{t}^*(c) := \frac{\sqrt{T}\,[\hat{\rho}^{-*}(c) - \hat{\rho}^{+*}(c)]}{\hat{\sigma}(c)}, \tag{5.1}$$

where $\hat{\rho}^{-*}(c)$ and $\hat{\rho}^{+*}(c)$ are constructed using the stationary bootstrap sample $\{(Y_t^*, X_t^*) : t = 1, \ldots, T\}$. Estimating the standard error $\hat{\sigma}(c)$, as discussed in 5.1, is crucial because it serves as the scale parameter of the above distribution. We propose using the nested resampling method to estimate the asymptotic variance $\sigma(c)$. For further details, see Hinkley and Shi (1989) and Tibshirani and Efron (1993).

Formally, we bootstrap a sequence of B_1 samples from the original data, from which we compute a sequence $\{\hat{\rho}_{(i)}^{-*}(c); \hat{\rho}_{(i)}^{+*}(c)\}_{i=1}^{B_1}$. The quantity $\hat{\sigma}^*(c)$ is then defined as the sample standard deviation of this sequence of B_1 values:

$$\hat{\sigma}^*(c) = \frac{1}{B_1 - 1}\sum_{i=1}^{B_1}[(\hat{\rho}_{(i)}^{-*}(c) - \hat{\rho}_{(i)}^{+*}(c)) - (\overline{\hat{\rho}_{(i)}^{-*}(c)} - \overline{\hat{\rho}_{(i)}^{+*}(c)})]^2. \tag{5.2}$$

To obtain the sampling distribution of the t-statistic $\hat{t}^*(c)$, we first generate bootstrap samples from the original data. For a given bootstrap sample j, we compute $\hat{\rho}(j)^{-*}(c)$ and $\hat{\rho}(j)^{+*}(c)$ using the equation in (2.3). The quantity $\hat{\sigma}_j^*(c)$ is estimated based on nested bootstrap samples with a sample size of B_1. Following Horowitz (2001), we adjust the t-statistic from the bootstrap samples to account for sampling bias:

$$\hat{t}_j^*(c) = \frac{[\hat{\rho}_j^{-*}(c) - \hat{\rho}_j^{+*}(c)] - [\hat{\rho}^{-*}(c) - \hat{\rho}^{+*}(c)]}{\hat{\sigma}_j^*(c))}. \tag{5.3}$$

We generate B bootstrap samples and estimate the empirical distribution F for $\{\hat{t}_j^*(c)\}_{j=1}^B$. We then report the percentile of $\hat{t}^*(c)$ under F. For a given significance level α, the null hypothesis of symmetric dependence will be rejected if $\hat{t}^*(c)$ falls in the upper $1-\alpha/2$ or lower $\alpha/2$ percentile of the empirical distribution F. For theoretical justification of this bootstrap approximation and the necessary assumptions, the reader is referred to Politis and Romano (1994*b*).

In order to generate data and assess the performance of the bootstrap-based test for symmetric dependence, we need to specify the marginal distributions of the random variables under study as well as the copulas that model the dependence between them in our simulation setting. For the latter, we consider copulas with varying levels of asymmetric dependence. The copulas under consideration include the Gaussian and Student's t copulas, which model symmetric exceedance dependence around their means, as well as the Clayton and Gumbel copulas, which capture asymmetric tail dependence. To evaluate the empirical size and power of the bootstrap-based test described in Theorem 2, we employ the following data generating processes (DGPs) based on mixed Gaussian-Clayton copulas; see Hong et al. (2007) for more details.

$$C_{mixt}(u,v,\rho,\tau,\kappa) = \kappa C_{nor}(u,v,\rho) + (1-\kappa)C_{clay}(u,v,\tau), \quad \kappa \in [0,1] \tag{5.4}$$

where κ is the mixture parameter. In other words, κ represents the weight assigned to the Gaussian copula. In our simulation design, we consider values of κ equal to 0, 0.25, 0.5, 0.75, and 1. By varying κ, we can achieve different levels of asymmetric dependence. Specifically, the mixture model nests the Gaussian copula as a special case when $\kappa = 1$ and reduces to the Clayton copula when $\kappa = 0$. The parameter ρ in the Gaussian copula corresponds to the correlation coefficient, while the parameter τ governs the dependence between the marginal distributions in the Clayton copula. A higher τ indicates stronger left-tail dependence.

For the individual distributions of the variables, we use marginal distributions based on a standard generalized autoregressive conditional heteroskedasticity (GARCH)(1, 1) process. The parameters for the copula-GARCH model used in our simulations were derived from real data. Specifically, we first fitted the copula-GARCH model to the equally-weighted return of the 5th smallest size portfolio and the value-weighted market returns [see the section on the empirical application for more information about these returns], then estimated the parameters of this model using maximum likelihood estimation (MLE). In the simulations, the data-generating processes (DGPs) are based on the copula-GARCH model with parameters set to the MLEs obtained from this fitting process. The MLEs for the Clayton copula parameter τ and the Gaussian copula parameter ρ are 4.351 and 0.914, respectively. The estimated GARCH parameters are presented in Table 1.

Table 1: Maximum likelihood estimates for GARCH(1, 1) processes

Parameter	Estimate	SE	Estimate	SE
μ_i	0.937	0.233	0.562	0.171
ω_i	4.977	2.321	1.139	0.556
α_i	0.137	0.046	0.107	0.029
β_i	0.730	0.094	0.844	0.036

Note: The table reports maximum likelihood estimates for parameters of the GARCH(1, 1) processes used to fit the equal-weighted return of the 5th smallest size portfolio and the value-weighted market return data. The GARCH models are used as the DGPs to simulate the data. The specification is set to follow a standard GARCH(1, 1) process: $r_{it} = \mu_i + \varepsilon_{i,t}$, where $\varepsilon_{i,t}$ is normally distributed with a time-varying variance $\sigma_{i,t}^2 = \omega_i + \alpha_i \varepsilon_{i,t-1}^2 + \beta_i \sigma_{i,t-1}^2$. μ_i is the unconditional mean of the variables. ω_i is the constant term in the time-varying conditional volatility process. α_i is the ARCH parameter and β_i is the GARCH parameter in the GARCH(1, 1) process. The sample period used for estimation is from January 1965 to December 2013.

Our nonparametric test relies on the bandwidth parameter k, which is necessary for estimating the Bernstein copula density (distribution). In our simulation, the optimal bandwidth parameter is selected by minimizing the mean squared error (MSE) of the copula density estimator, following the approach outlined by Rose (2015). A practical bandwidth can also be determined using an approach similar to that proposed by Omelka, Gijbels and Veraverbeke (2009) for kernel-based copula estimation. However, this alternative method is not explored in the current paper and is left for future research. Omelka et al. (2009)'s method involves an Edgeworth expansion of the asymptotic distribution of the test statistics.

The simulations are conducted using 1000 replications of copula-GARCH samples, where the dependence structure adheres to the mixed copula described in Equation 5.4. Specifically, we calculate the empirical size and power of our proposed test using different sample sizes (T = 240, 420, 600, and 1200). The empirical size and power are computed as the relative frequency of rejecting the null hypothesis of symmetric dependence at different exceedance levels $c = 0, 1, 1.5$ in the simulated samples.

Tables 2, 3, and 4 present the results for the empirical size and power of our asymmetric dependence test at the nominal size levels of 10%, 5%, and 1% for different exceedance levels $c = 0, 1, 1.5$.

From the above results and when $c = 0$ (refer to Table 2), we note the following. When $\kappa = 1$, i.e., equation 5.4 reduces to the Gaussian copula with symmetric tail dependence. In this case, our proposed test demonstrates accurate sizes in finite samples, showing robustness across different sample sizes. When the data exhibits minimal deviation from normality with $\kappa = 0.75$, the rejection rates remain low when $T = 240$. This low rejection rate is attributed to the relatively smaller number of samples available at higher exceedance levels. However, the power of the test increases significantly as κ decreases to 0.5, 0.25, and 0.

Furthermore, a significant improvement in the power of our test is observed as the sample size increases. Notably, when the data-generating processes (DGPs) exhibit stronger asymmetric dependence ($\kappa = 0.25$ and 0), the performance of our test further improves, with its power converging to 1 at a faster rate.

For the other exceedance levels $c = 1$ and $c = 1.5$ (refer to Tables 3 and 4), our test continues to show accurate empirical sizes for all nominal levels of 10%, 5%, and 1%. Regarding empirical power, the pattern indicates that a larger sample size is required for our test to achieve good empirical power when exceedance levels in the farther tails are considered.

Table 2: Empirical Size and power of the test of symmetric dependence test, with $c = 0$

Sample size (T)	Nominal size	100% (Size)	75%	50%	25%	0%
240	10%	0.091	0.448	0.861	0.994	1.000
	5%	0.042	0.309	0.758	0.986	0.997
	1%	0.009	0.095	0.495	0.884	0.988
420	10%	0.121	0.511	0.927	1.000	1.000
	5%	0.066	0.440	0.936	1.000	1.000
	1%	0.014	0.240	0.802	1.000	1.000
600	10%	0.112	0.755	0.996	1.000	1.000
	5%	0.063	0.672	0.991	1.000	1.000
	1%	0.012	0.335	0.950	1.000	1.000
1200	10%	0.120	0.994	1.000	1.000	1.000
	5%	0.051	0.991	1.000	1.000	1.000
	1%	0.012	0.917	1.000	1.000	1.000

Note: The table reports empirical size and power of the test of symmetric dependence. The exceedance level c is set to 0 in all cases. The inference is based on 199 stationary bootstrap resamplings and 1000 replications. We construct the sampling distribution for our proposed test using the pivotal bootstrap resampling approach. We employ a stationary block bootstrap method to take into account the dependent structure in weakly dependent time series data. Stationarity is ensured by letting the length of each block be randomly sampled from the geometric distribution; see Politis and White (2004).

Table 3: Empirical Size and power of the test of symmetric dependence test, with $c = 1$

Sample size (T)	Nominal size	100% (Size)	75%	50%	25%	0%
240	10%	0.107	0.261	0.612	0.871	0.981
	5%	0.048	0.171	0.445	0.791	0.962
	1%	0.008	0.053	0.210	0.533	0.791
420	10%	0.120	0.350	0.791	0.980	1.000
	5%	0.051	0.241	0.682	0.945	0.994
	1%	0.010	0.069	0.510	0.811	0.959
600	10%	0.092	0.461	0.931	0.992	1.000
	5%	0.050	0.311	0.854	0.985	1.000
	1%	0.009	0.119	0.510	0.941	1.000
1200	10%	0.102	0.750	0.938	1.000	1.000
	5%	0.051	0.610	0.981	1.000	1.000
	1%	0.003	0.321	0.911	1.000	1.000
1800	10%	0.101	0.901	1.000	1.000	1.000
	5%	0.051	0.890	1.000	1.000	1.000
	1%	0.010	0.512	1.000	1.000	1.000

Note: The table reports empirical size and power of the test of symmetric dependence. The exceedance level c is set to 1 in all cases. The inference is based on 199 stationary bootstrap resamplings and 1000 replications. We construct the sampling distribution for our proposed test using the pivotal bootstrap resampling approach. We employ a stationary block bootstrap method to take into account the dependent structure in weakly dependent time series data. Stationarity is ensured by letting the length of each block be randomly sampled from the geometric distribution; see Politis and White (2004).

Table 4: Empirical Size and power of the test of symmetric dependence test, with $c = 1.5$

Sample size (T)	Nominal size	100% (Size)	75%	50%	25%	0%
240	10%	0.110	0.155	0.360	0.560	0.786
	5%	0.040	0.082	0.221	0.391	0.634
	1%	0.006	0.015	0.071	0.160	0.321
420	10%	0.110	0.180	0.501	0.710	0.958
	5%	0.052	0.111	0.371	0.689	0.932
	1%	0.012	0.032	0.157	0.417	0.717
600	10%	0.091	0.241	0.641	0.921	0.995
	5%	0.045	0.142	0.505	0.831	0.920
	1%	0.011	0.027	0.211	0.603	0.821
1200	10%	0.112	0.371	0.834	0.959	1.000
	5%	0.053	0.314	0.788	0.968	1.000
	1%	0.013	0.121	0.491	0.946	0.999
1800	10%	0.100	0.577	0.927	1.000	1.000
	5%	0.048	0.471	0.944	1.000	1.000
	1%	0.010	0.236	0.903	1.000	1.000
2400	10%	0.100	0.693	0.995	1.000	1.000
	5%	0.047	0.572	0.981	1.000	1.000
	1%	0.011	0.272	0.914	1.000	1.000
3600	10%	0.103	0.901	1.000	1.000	1.000
	5%	0.044	0.821	1.000	1.000	1.000
	1%	0.008	0.520	1.000	1.000	1.000

Note: The table reports empirical size and power of the test of symmetric dependence. The exceedance level c is set to 1.5 in all cases. The inference is based on 199 stationary bootstrap resamplings and 1000 replications. We construct the sampling distribution for our proposed test using the pivotal bootstrap resampling approach. We employ a stationary block bootstrap method to take into account the dependent structure in weakly dependent time series data. Stationarity is ensured by letting the length of each block be randomly sampled from the geometric distribution; see Politis and White (2004).

We conduct additional simulations to evaluate the performance of our proposed test when the GARCH error terms follow a non-normal distribution. Specifically, we use the same simulation setup as before, but this time we generate the marginal distribution of each variable based on two alternative distributions: the Student t and the skewed t distributions.

Tables 5 and 6 present the test performance when the marginal distributions follow the t and skewed t distributions, respectively. In Table 5, we observe that the empirical size is well maintained, even for the smallest sample size, $T = 240$. The empirical powers are also quite similar to those in the benchmark case (normal errors) and generally converge to 1 as the sample size increases to

$T = 1,200$. In Table 6, the empirical powers are slightly lower than in the benchmark case but still converge to 1 when the sample size is large.

Overall, we conclude that our asymmetric dependence test performs very well in finite samples, even when the marginal distributions exhibit a degree of leptokurtic behavior.

Table 5: Empirical Size and power of the test of symmetric dependence test, t marginal distributions ($c = 0$)

Sample size (T)	Nominal size	100% (Size)	75%	50%	25%	0%
240	10%	0.119	0.391	0.841	0.981	0.986
	5%	0.058	0.311	0.746	0.941	0.985
	1%	0.009	0.098	0.467	0.801	0.942
420	10%	0.119	0.534	0.932	0.971	0.989
	5%	0.053	0.469	0.931	0.973	0.992
	1%	0.012	0.271	0.801	0.971	0.987
600	10%	0.106	0.742	0.986	0.988	0.991
	5%	0.055	0.641	0.972	0.971	0.991
	1%	0.011	0.317	0.899	0.962	0.979
1200	10%	0.115	0.949	0.989	1.000	1.000
	5%	0.055	0.919	1.000	0.998	1.000
	1%	0.013	0.818	0.991	0.989	0.998

Note: The table reports empirical size and power of the test of symmetric dependence for t marginal distributions. The nominal sizes are set to 10%, 5%, and 1%, respectively. All random samples are generated by the mixture copula in Equation 5.4. The inference is based on 199 stationary bootstrap resamplings and 1000 replications, and the exceedance level c is set to 0 in all cases.

Table 6: Empirical Size and power of the test of symmetric dependence test, with skewed t marginal distributions ($c = 0$)

Sample size (T)	Nominal size	100% (Size)	75%	50%	25%	0%
240	10%	0.119	0.312	0.789	0.957	0.989
	5%	0.094	0.193	0.651	0.935	0.977
	1%	0.015	0.058	0.381	0.781	0.928
420	10%	0.172	0.312	0.812	0.912	0.989
	5%	0.161	0.251	0.816	0.972	0.976
	1%	0.021	0.097	0.602	0.891	0.960
600	10%	0.271	0.471	0.921	0.992	0.995
	5%	0.153	0.351	0.872	0.913	0.971
	1%	0.051	0.135	0.712	0.989	0.993
1200	10%	0.403	0.711	0.995	0.993	0.997
	5%	0.301	0.568	0.989	0.991	0.996
	1%	0.108	0.291	0.958	0.983	0.996

Note: The table reports empirical size and power of the test of symmetric dependence for skewed t marginal distributions. The nominal sizes are set to 10%, 5%, and 1%, respectively. All random samples are generated by the mixture copula in Equation 5.4. The inference is based on 199 stationary bootstrap resamplings and 1000 replications, and the exceedance level c is set to 0 in all cases.

5.2 Bootstrap-Based Bias Correction

Here, we provide additional simulation results to examine a bootstrap bias-corrected estimator for measures of extreme dependence. Our estimators are again motivated by the stationary block bootstrap method, which accounts for the dependent structure in weakly dependent time series data. The procedure is straightforward: we first use the bootstrapped sample to estimate the finite sample bias in the nonparametric estimators of Bernstein copula-based measures of extreme dependence. We then subtract this bias term to obtain the bootstrap bias-corrected estimates. The estimates can be obtained easily through the following four steps:

Step 1 We draw a stationary bootstrap sample $\{(X_t^\star, Y_t^\star)\}_{t=1}^T$;

Step 2 Based on the sample $\{(X_t^\star, Y_t^\star)\}_{t=1}^T$, we compute the bootstrap estimators of the measures of extreme dependence;

Step 3 We repeat steps 1 and 2 B times so that we get $\hat{\rho}^{-*}(c)$ and $\hat{\rho}_j^{+*}(c)$ for $j = 1, \ldots, B$; and

Step 4 We approximate the bias term $Bias^+ = E[\hat{\rho}^+(c)] - \rho^+(c)$ by the corresponding bootstrapped Bias$^{+\star}$= $E^\star[\hat{\rho}^{+\star}(c)] - \hat{\rho}^+(c)$, where $E^\star$ is the expectation based on the bootstrapped distribution of $\hat{\rho}^{+\star}(c)$ and $\hat{\rho}^+(c)$ is the estimate of $\rho^+(c)$ using the original sample. This suggests the bias estimate

$$\hat{Bias}^{+\star} = \frac{1}{B}\sum_{j=1}^{B} \hat{\rho}_j^{+*}(c) - \hat{\rho}^{+\star}(c)).$$

Hence, a bootstrap bias-corrected estimators of measures of extreme dependence can be defined as follows:

$$\hat{\rho}_{BC}^{+\star}(c) = \hat{\rho}^{+}(c) - \hat{Bias}^{+\star}$$

and

$$\hat{\rho}_{BC}^{-\star}(c) = \hat{\rho}^{-}(c) - \hat{Bias}^{-\star}.$$

Tables 7-9 present the simulation results for the bootstrap bias-corrected estimates of measures of extreme dependence with $c = 0$. We use the same data-generating processes (DGPs) as described in Subsection 1. The bias terms and the average values of the bootstrap bias-corrected estimates are computed based on 1000 simulations, with $B = 199$ bootstrap replications.

From the three tables, particularly Table 7, where these measures are expected to be exactly zero under the Gaussian copula, we observe that the bootstrap bias-corrected estimators perform better than the bias-uncorrected estimators.

Table 7: Simulation results for Bias correction under Gaussian copula, $c = 0$

	$\hat{\rho}^{-}(c)$			$\hat{\rho}^{+}(c)$		
	$T = 100$	$T = 150$	$T = 200$	$T = 100$	$T = 150$	$T = 200$
Bias-Uncorrected Estimator	0.11	0.10	0.09	0.11	0.10	0.09
Bias-Corrected Estimator	0.006	0.001	0.0004	0.006	0.001	0.0004

Table 8: Simulation results for Bias correction under mixture (50%-50%) Clayton-Gaussian copula, $c = 0$

	$\hat{\rho}^{-}(c)$			$\hat{\rho}^{+}(c)$		
	$T = 100$	$T = 150$	$T = 200$	$T = 100$	$T = 150$	$T = 200$
Bias-Uncorrected Estimator	0.12	0.131	0.141	0.052	0.061	0.061
Bias-Corrected Estimator	0.051	0.061	0.012	0.006	0.016	0.025

Table 9: Simulation results for Bias correction under Clayton copula, $c = 0$

	$\hat{\rho}^-(c)$			$\hat{\rho}^+(c)$		
	$T = 100$	$T = 150$	$T = 200$	$T = 100$	$T = 150$	$T = 200$
Bias-Uncorrected Estimator	0.13	0.15	0.16	0.025	0.018	0.017
Bias-Corrected Estimator	0.054	0.061	0.063	0.012	0.012	0.016

6 Empirical application

In this section, we apply the non-parametric test proposed in the previous sections to commonly used U.S. and international equity portfolios. Our goal is to investigate whether asymmetric comovement is a universal and prevalent phenomenon in stock returns. We compare the results of our test (DST) to those obtained from the Jiang, Maasoumi, Pan and Wu (2018) test (jmpw) by directly comparing the computed p-values for testing the symmetric dependence in the U.S. and other international financial markets.

Our analysis focuses on the excess returns of value-weighted size and book-to-market decile portfolios, as well as equal-weighted decile momentum portfolios in the U.S. market. The dataset consists of monthly U.S. T-bill observations from January 1965 to December 2016, totaling 624 observations.[1]

Additionally, we examine the asymmetric comovement of portfolio returns in several other countries: Canada, France, Germany, Japan, Switzerland, and the UK. Specifically, we consider portfolios sorted by book-to-market (B/M), earnings-to-price (E/P), and cash earnings-to-price (CE/P), and investigate their return comovement with the market return of their respective countries. Finally, for each country, we test whether their market returns are asymmetrically dependent on the U.S. stock market return.

As discussed earlier, the implementation of our non-parametric approach to test for asymmetric dependence relies on the Bernstein copula density. First, we provide estimates of the measures of extreme tail dependence, ρ^+ and ρ^-. Second, we present the p-values of our proposed copula-based asymmetric dependence test across all three sets of portfolios, which are sorted by size, book-to-market ratio, and past cumulative returns. These p-values are calculated using 399 stationary bootstrap resamplings, with the exceedance level set to $c = 0$. The empirical results are summarized in Tables 10 and 11 of Appendix A.

Table 10 in Appendix A shows that our nonparametric test rejects the symmetry dependence hy-

[1]The data are available on Kenneth French's website. Note that the international portfolio data spans from January 1975 to December 2016, except for Canada, which is available from January 1977 to December 2016.

pothesis for the first eight smallest-size portfolios, indicating that larger firms exhibit more symmetric comovement with the market. For value-weighted book-to-market portfolios, we reject the null hypothesis of symmetric dependence for the tenth BE/ME portfolio at the 10% significance level. These findings align with previous literature, which shows that value stocks tend to have more asymmetric comovement with the market; see Ang and Chen (2002), Jondeau (2016), and Jiang, Maasoumi, Pan and Wu (2018).

Among momentum-sorted portfolios, our test also rejects the null hypothesis of symmetric dependence. In Table 11 of of Appendix A, we observe asymmetry across all three sets of portfolios sorted by size, book-to-market ratio, and past cumulative returns in the markets of Canada, France, Germany, and Switzerland. Concerning dependence on the U.S. market, our proposed test indicates a higher dependence during market downturns between the U.S. market and these international markets.

7 Conclusion

We introduced measures of extreme dependence between random variables by defining a Kullback-Leibler relative entropy in terms of copula densities. To estimate these measures consistently in a nonparametric way, we employed a Bernstein copula density estimator and derived Bahadur-type representations for these estimators. We then established the asymptotic distribution of a test for symmetric dependence, which was constructed using the aforementioned measures. We evaluated the properties of this test under both global and local alternatives. Additionally, we demonstrated the validity of a bootstrap-based test for symmetric dependence, suitable for finite-sample contexts.

A Monte Carlo simulation study indicated that the bootstrap-based test demonstrates valid size and strong power across various data-generating processes and sample sizes, offering a reliable approximation of the asymptotic-based test in finite-sample contexts. Finally, we presented an empirical application that highlights the practical utility of these extreme dependence measures, specifically quantifying the degree of extreme dependence between the U.S. financial market and various developed and emerging financial markets.

References

Ang, A. and Bekaert, G. (2002), ‘International asset allocation with regime shifts’, *The review of financial studies* **15**(4), 1137–1187.

Ang, A. and Chen, J. (2002), ‘Asymmetric correlations of equity portfolios’, *Journal of financial Economics* **63**(3), 443–494.

Belalia, M., Bouezmarni, T., Lemyre, F. and Taamouti, A. (2017), ‘Testing independence based on bernstein empirical copula and copula density’, *Journal of Nonparametric Statistics* **29**(2), 346–380.

Bouezmarni, T., Rombouts, J. V. and Taamouti, A. (2010), ‘Asymptotic properties of the bernstein density copula estimator for α-mixing data’, *Journal of Multivariate Analysis* **101**(1), 1–10.

Bouezmarni, T., Rombouts, J. V. and Taamouti, A. (2012), ‘A nonparametric copula based test for conditional independence with applications to granger causality’, *Journal of Business Economic Statistics* pp. 275–287.

Doukhan, P. and Wintenberger, O. (2008), ‘Weakly dependent chains with infinite memory’, *Stochastic Processes and their Applications* **118**(11), 1997–2013.

Hall, P. (1984), ‘Central limit theorem for integrated square error of multivariate nonparametric density estimators’, *Journal of multivariate analysis* **14**(1), 1–16.

Hinkley, D. and Shi, S. (1989), ‘Importance sampling and the nested bootstrap’, *Biometrika* **76**(3), 435–446.

Hong, Y., Tu, J. and Zhou, G. (2007), ‘Asymmetries in stock returns: Statistical tests and economic evaluation’, *The Review of Financial Studies* **20**(5), 1547–1581.

Horowitz, J. L. (2001), The bootstrap, *in* ‘Handbook of econometrics’, Vol. 5, Elsevier, pp. 3159–3228.

Hwang, E. and Shin, D. W. (2011), ‘Stationary bootstrapping for non-parametric estimator of nonlinear autoregressive model’, *Journal of Time Series Analysis* **32**(3), 292–303.

Janssen, P., Swanepoel, J. and Veraverbeke, N. (2014), ‘A note on the asymptotic behavior of the bernstein estimator of the copula density’, *Journal of Multivariate Analysis* **124**, 480–487.

Jiang, L., Maasoumi, E., Pan, J. and Wu, K. (2018), ‘A test of general asymmetric dependence’, *Journal of Applied Econometrics* **33**(7), 1026–1043.

Jiang, L., Wu, K. and Zhou, G. (2018), ‘Asymmetry in stock comovements: An entropy approach’, *Journal of Financial and Quantitative Analysis* **53**(4), 1479–1507.

Jondeau, E. (2016), ‘Asymmetry in tail dependence in equity portfolios’, *Computational Statistics & Data Analysis* **100**, 351–368.

Kullback, S. and Leibler, R. A. (1951), ‘On information and sufficiency’, *The annals of mathematical statistics* **22**(1), 79–86.

Kunsch, H. R. (1989), ‘The jackknife and the bootstrap for general stationary observations’, *The annals of Statistics* pp. 1217–1241.

Liu, W. and Wu, W. B. (2010), ‘Simultaneous nonparametric inference of time series1’, *The Annals of Statistics* **38**(4), 2388–2421.

Longin, F. and Solnik, B. (2001), ‘Extreme correlation of international equity markets’, *The journal of finance* **56**(2), 649–676.

Omelka, M., Gijbels, I. and Veraverbeke, N. (2009), ‘Improved kernel estimation of copulas: weak convergence and goodness-of-fit testing’, *The Annals of Statistics* **37**(5B), 3023–3058.

Patton, A. J. (2004), ‘On the out-of-sample importance of skewness and asymmetric dependence for asset allocation’, *Journal of financial econometrics* **2**(1), 130–168.

Patton, A., Politis, D. N. and White, H. (2009), ‘Correction to automatic block-length selection for the dependent bootstrap by d. politis and h. white’, *Econometric Reviews* **28**(4), 372–375.

Politis, D. N. and Romano, J. P. (1994*a*), ‘Limit theorems for weakly dependent hilbert space valued random variables with application to the stationary bootstrap’, *Statistica Sinica* pp. 461–476.

Politis, D. N. and Romano, J. P. (1994*b*), ‘The stationary bootstrap’, *Journal of the American Statistical association* **89**(428), 1303–1313.

Politis, D. N. and White, H. (2004), ‘Automatic block-length selection for the dependent bootstrap’, *Econometric Reviews* **23**(1), 53–70.

Rose, D. (2015), Modeling and estimating multivariate dependence structures with the Bernstein copula, PhD thesis, lmu.

Sancetta, A. and Satchell, S. (2004), 'The bernstein copula and its applications to modeling and approximations of multivariate distributions', *Econometric theory* **20**(3), 535–562.

Sklar, M. (1959), Fonctions de répartition à n dimensions et leurs marges, *in* 'Annales de l'ISUP', Vol. 8, pp. 229–231.

Song, X. and Taamouti, A. (2017), 'Measuring nonlinear granger causality in quantiles', *Journal of Business Economic Statistics* .

Tibshirani, R. J. and Efron, B. (1993), 'An introduction to the bootstrap', *Monographs on statistics and applied probability* **57**(1).

Appendix A: Empirical results

This appendix presents the tables of empirical results discussed in the main text. The tables include detailed outputs from our nonparamatric tests of symmetric dependence in the U.S. and other international financial markets, providing insights and supporting evidence for the findings reported in the paper.

Table A.10: Measuring dependence in the U.S. stock portfolios

Value-weighted size portfolios					
Portfolio	$\hat{\rho}^{-}(c)$	$\hat{\rho}^{+}(c)$	$\Delta\rho$	DST	$jmpw$
size1	0.30	0.21	0.09	0.00	0.00
size2	0.37	0.29	0.08	0.00	0.00
size3	0.41	0.35	0.06	0.01	0.01
size4	0.42	0.37	0.05	0.02	0.03
size5	0.46	0.43	0.03	0.10	0.11
size6	0.51	0.48	0.03	0.12	0.13
size7	0.56	0.55	0.01	0.37	0.39
size8	0.62	0.60	0.02	0.38	0.36
size9	0.66	0.69	0.03	0.17	0.15
size10	0.65	0.70	0.05	0.09	0.08
Value-weighted book-to-market portfolios					
Portfolio	$\hat{\rho}^{-}(c)$	$\hat{\rho}^{+}(c)$	$\Delta\rho$	DTS	$jmpw$
BE/ME 1	0.45	0.45	0.00	0.93	0.91
BE/ME 2	0.55	0.56	0.01	0.89	0.86
BE/ME 3	0.56	0.52	0.04	0.13	0.14
BE/ME 4	0.50	0.51	0.01	0.98	0.96
BE/ME 5	0.47	0.44	0.03	0.12	0.10
BE/ME 6	0.44	0.42	0.02	0.54	0.55
BE/ME 7	0.38	0.38	0.00	0.98	0.99
BE/ME 8	0.38	0.36	0.02	0.33	0.35
BE/ME 9	0.40	0.37	0.03	0.35	0.32
BE/ME 10	0.32	0.28	0.04	0.05	0.06
Equal-weighted momentum portfolios					
Portfolio	$\hat{\rho}^{-}(c)$	$\hat{\rho}^{+}(c)$	$\Delta\rho$	DST	$jmpw$
L	0.26	0.18	0.08	0.00	0.00
2	0.33	0.26	0.07	0.00	0.00
3	0.36	0.30	0.06	0.00	0.00
4	0.38	0.33	0.05	0.00	0.00
5	0.38	0.33	0.05	0.00	0.02
6	0.40	0.35	0.05	0.00	0.02
7	0.41	0.35	0.06	0.00	0.01
8	0.39	0.32	0.07	0.00	0.00
9	0.39	0.29	0.10	0.00	0.00
W	0.33	0.26	0.07	0.00	0.00

Note: The table reports the estimates and p-values (for the statistical significance) of exceedance dependence measures for common U.S. portfolios sorted by size, book-to-market, and momentum at the exceedance level $c = 0$.

Table A.11: Testing asymmetric dependence between U.S. and international stock portfolios

International stock portfolios														
	BE/ME sorted portfolios				E/P sorted portfolios				CE/P sorted portfolios				Comove with US	
	High		Low		High		Low		High		Low			
Country	DST	jmpw	DST	jmpw	DST	jmpw	DST	jmpw	DST	jmpw	DST	jmpw	DST	jmpw
Canada	0.25	0.20	0.18	0.19	0.25	0.26	0.02	0.04	0.02	0.03	0.02	0.03	0.00	0.07
France	0.05	0.07	0.61	0.59	0.02	0.01	0.30	0.33	0.05	0.06	0.30	0.34	0.00	0.00
Germ.	0.01	0.01	0.03	0.05	0.27	0.29	0.20	0.18	0.16	0.17	0.39	0.42	0.00	0.00
Japan	0.28	0.31	0.50	0.48	0.61	0.67	0.72	0.76	0.50	0.52	0.78	0.81	0.00	0.00
Switz.	0.52	0.62	0.48	0.45	0.10	0.12	0.05	0.07	0.25	0.30	0.60	0.65	0.00	0.00
UK	0.73	0.70	0.80	0.78	0.55	0.58	0.80	0.82	0.90	0.96	0.65	0.68	0.00	0.00

Note: This table reports the p-values for testing dependence between U.S. market and international stock portfolios sorted by book-to-market (B/M), earnings to price (E/P), and cash earnings to price (CE/P) at the exceedance level $c = 0$. The p-values are computed based on 399 stationary bootstrap resamplings.

Appendix B: Proofs of the main theoretical results

This appendix provides the detailed proofs of the main theoretical results presented in the paper.

Proof of Theorem 1: The following proof is for the bivariate case $d_X + d_Y = 2$. For the more general multidimensional case, $d_X + d_Y \geq 3$, the proof can be obtained in a similar way with slightly more complex notations. We will also focus on the asymptotic behavior of the estimator $\hat{\rho}^-(c)$, as the asymptotic behavior of estimator $\hat{\rho}^+(c)$ can be studied analogously.

First of all, denote $F_{XYT}(x,y) = T^{-1}\sum_{t=1}^{T} \mathscr{I}(X_t \leq x, Y_t \leq y)$ to be the empirical distribution function of the sample $\left\{(X_t, Y_t)'\right\}_{t=1}^{T}$. Recall that the estimator $\hat{\rho}^-(c)$ can be written as:

$$
\begin{aligned}
\hat{\rho}^-(c) &= \frac{1}{T}\sum_{t=1}^{T} \log\left[c_{k,T}(\mathbf{V_t})\right] \mathscr{I}(X_t \leq -c, Y_t \leq -c) \\
&= \int \log\left[c_{k,T}\left(F_{XT}(x), F_{YT}(y)\right)\right] \mathscr{I}(x \leq -c, y \leq -c)\, dF_{XYT}(x,y) \\
&= \int \log\left[c_{k,T}\left(F_{XT}(x), F_{YT}(y)\right)\right] \mathscr{I}(x \leq -c, y \leq -c)\, d\left(F_{XYT}(x,y) - F_{XY}(x,y)\right) \\
&\quad + \int \log\left[c_{k,T}\left(F_{XT}(x), F_{YT}(y)\right)\right] \mathscr{I}(x \leq -c, y \leq -c)\, dF_{XY}(x,y),
\end{aligned}
$$

Thus,

$$
\begin{aligned}
\hat{\rho}^-(c) - \rho^-(c) &= \int \log\left[c_{k,T}\left(F_{XT}(x), F_{YT}(y)\right)\right] \mathscr{I}(x \leq -c, y \leq -c)\, d\left(F_{XYT}(x,y) - F_{XY}(x,y)\right) \\
&\quad + \int \log\left[\frac{c_{k,T}\left(F_{XT}(x), F_{YT}(y)\right)}{c\left(F_X(x), F_Y(y)\right)}\right] \mathscr{I}(x \leq -c, y \leq -c)\, dF_{XY}(x,y) \\
&: = A_{1T}(c) + A_{2T}(c).
\end{aligned}
$$

Hereafter, we shall deal with the terms $A_{1T}(c)$ and $A_{2T}(c)$ separately. Before proceeding, recall the following generalized errors:

$$
\varepsilon_{1t}^-(c) = \log\left[c\left(F_X(X_t), F_Y(Y_t)\right)\right] \mathscr{I}(X_t \leq -c, Y_t \leq -c) - \rho^-(c),
$$

and

$$
\begin{aligned}
\varepsilon_{2t}^-(c) =& \mathscr{I}(X_t \leq -c, Y_t \leq -c) - C\left(F_X(-c), F_Y(-c)\right) \\
& - C_{u_1}\left(F_X(-c), F_Y(-c)\right)\left[\mathscr{I}(X_t \leq -c) - F_X(-c)\right] \\
& - C_{u_2}\left(F_X(-c), F_Y(-c)\right)\left[\mathscr{I}(Y_t \leq -c) - F_Y(-c)\right],
\end{aligned}
$$

where $C(F_X(-c), F_Y(-c)) = E[\mathscr{I}(X_t \leq -c, Y_t \leq -c)]$ and $C_{u_l}(u_1, u_2) := \partial C(u_1, u_2)/\partial u_l$ for $l = 1, 2$. It is straightforward to see that under mild regularity conditions both generalized errors $\varepsilon_{1t}^{-}(c)$ and $\varepsilon_{2t}^{-}(c)$ have mean zero and finite variance.

As we will show in the following, it turns out that the asymptotic behavior of $\sqrt{T}[\hat{\rho}^{-}(c) - \rho^{-}(c)]$ is determined by the following two results:

$$\textbf{Result (i)}: \sqrt{T} A_{1T}(c) = \frac{1}{\sqrt{T}} \sum_{t=1}^{T} \varepsilon_{1t}^{-}(c) + o_p(1),$$

and

$$\textbf{Result (ii)}: \sqrt{T} A_{2T}(c) = \frac{1}{\sqrt{T}} \sum_{t=1}^{T} \varepsilon_{2t}^{-}(c) + o_p(1).$$

Therefore, results (i) and (ii) together imply immediately that $\sqrt{T}[\hat{\rho}^{-}(c) - \rho^{-}(c)]$ admits an asymptotic Bahadur representation

$$\frac{1}{\sqrt{T}} \sum_{t=1}^{T} \left\{\varepsilon_{1t}^{-}(c) + \varepsilon_{2t}^{-}(c)\right\} + o_p(1),$$

as provided in Theorem 1. To save space, henceforth,we omit the dependence of those quantities (e.g. $A_{1T}(c)$ and $\varepsilon_{1t}(c)$) on the exceedance level c.

Proof of Result (i): First notice the following decomposition:

$$\begin{aligned} A_{1T} =& \int \log\left[c\left(F_{XT}(x), F_{YT}(y)\right)\right] \mathscr{I}(x \leq -c, y \leq -c)\, d\left(F_{XYT}(x,y) - F_{XY}(x,y)\right) \\ &+ \int \log\left[\frac{c_{k,T}\left(F_{XT}(x), F_{YT}(y)\right)}{c\left(F_{XT}(x), F_{YT}(y)\right)}\right] \mathscr{I}(x \leq -c, y \leq -c)\, d\left(F_{XYT}(x,y) - F_{XY}(x,y)\right) \\ =& A_{11T} + A_{12T}, \end{aligned}$$

First of all, recall the following theorems: (1) Glivenko-Cantelli theorem for the stationary ergodic processes $\{X_t\}$ and $\{Y_t\}$ that states that the empirical distribution functions $F_{XT}(x)$ and $F_{YT}(y)$ of X and Y converge almost surely to $F_X(x)$ and $F_Y(y)$, respectively, over $x \in \mathbb{R}$ and $y \in \mathbb{R}$ uniformly and (2) Donsker weak invariance principle for weakly dependent stationary process $(X_t, Y_t)'$ that states that $\sqrt{T}(F_{XYT}(x,y) - F_{XY}(x,y)) \Rightarrow G_\infty$, where G_∞ is a Gaussian process with zero mean and covariance structure given by $\Omega(x_1, y_1; x_2, y_2)$ [see e.g. Theorem 2.1 in Doukhan and Wintenberger (2008)], such that the supremum of absolute value, i.e. $\sup_{(x,y)} \sqrt{T}|F_{XYT}(x,y) - F_{XY}(x,y)|$ converges in distribution to the law of the same functional of the Gaussian process G_∞, i.e. $\sup_{(x,y)} |G_\infty(x,y)|$, and thus $\sup_{(x,y)} \sqrt{T}|F_{XYT}(x,y) - F_{XY}(x,y)| = O_p(1)$. Based on these classical results, it is straightforward to show that

$$\begin{aligned}
\sqrt{T}A_{11T} &= \int \log\left[c\left(F_X(x),F_Y(y)\right)\right]\mathscr{I}\left(x\le -c, y\le -c\right)d\sqrt{T}\left(F_{XYT}(x,y)-F_{XY}(x,y)\right) \\
&\quad + \int \log\left[\frac{c\left(F_{XT}(x),F_{YT}(y)\right)}{c\left(F_X(x),F_Y(y)\right)}\right]\mathscr{I}\left(x\le -c, y\le -c\right)d\sqrt{T}\left(F_{XYT}(x,y)-F_{XY}(x,y)\right) \\
&= \int \log\left[c\left(F_X(x),F_Y(y)\right)\right]\mathscr{I}\left(x\le -c, y\le -c\right)d\sqrt{T}\left(F_{XYT}(x,y)-F_{XY}(x,y)\right)+o_p(1) \\
&= \frac{1}{\sqrt{T}}\sum_{t=1}^{T}\{\log\left[c\left(F_X(X_t),F_Y(Y_t)\right)\right]\mathscr{I}\left(X_t\le -c, Y_t\le -c\right)-E\left[\log\left[c\left(F_X(X_t),F_Y(Y_t)\right)\right]\mathscr{I}\left(X_t\le -c, Y_t \le\right.\right. \\
&= \frac{1}{\sqrt{T}}\sum_{t=1}^{T}\varepsilon_{1t}^{-}+o_p(1),
\end{aligned}$$

where the first step follows from observing that

$$\begin{aligned}
&\left|\int \log\left[\frac{c\left(F_{XT}(x),F_{YT}(y)\right)}{c\left(F_X(x),F_Y(y)\right)}\right]\mathscr{I}\left(x\le -c, y\le -c\right)d\sqrt{T}\left(F_{XYT}(x,y)-F_{XY}(x,y)\right)\right| \\
\le\ & C\sup_{(x,y)}\left|\frac{c(F_{XT}(x),F_{YT}(y))-c(F_X(x),F_Y(y))}{c(F_X(x),F_Y(y))}\right|\int_{-\infty}^{-c}d\sup_{(x,y)}\sqrt{T}\left|F_{XYT}(x,y)-F_{XY}(x,y)\right| \\
=\ & o_p(1),
\end{aligned}$$

using $\sup_{(x,y)}|c(F_{XT}(x),F_{YT}(y))-c(F_X(x),F_Y(y))|\le C\left(\sup_x|F_{XT}(x)-F_X(x)|+\sup_y|F_{XT}(y)-F_X(y)|\right)=o_p(1)$ and taking into account the fact that the trajectories of the limiting Gaussian process G_∞ are bounded and continuous almost surely. It remains to show that $\sqrt{T}A_{12T}=o_p(1)$, which is can be proved as follows:

$$\begin{aligned}
&\left|\int \log\left[\frac{c_{k,T}\left(F_{XT}(x),F_{YT}(y)\right)}{c\left(F_{XT}(x),F_{YT}(y)\right)}\right]\mathscr{I}\left(x\le -c, y\le -c\right)d\sqrt{T}\left(F_{XYT}(x,y)-F_{XY}(x,y)\right)\right| \\
&\le C\sup_{(x,y)}\left|\frac{c_{k,T}(F_{XT}(x),F_{YT}(y))-c(F_{XT}(x),F_{YT}(y))}{c(F_{XT}(x),F_{YT}(y))}\right|\int_{-\infty}^{-c}d\sup_{(x,y)}\sqrt{T}\left|F_{XYT}(x,y)-F_{XY}(x,y)\right| \\
&\le C\left(\sup_{(x,y)}\left|\frac{c_{k,T}(F_{XT}(x),F_{YT}(y))-c_{k,T}(F_X(x),F_Y(y))}{c(F_{XT}(x),F_{YT}(y))}\right|+\sup_{(x,y)}\left|\frac{c_{k,T}(F_X(x),F_Y(y))-c(F_X(x),F_Y(y))}{c(F_{XT}(x),F_{YT}(y))}\right|\right. \\
&\quad\left.+\sup_{(x,y)}\left|\frac{c(F_{XT}(x),F_{YT}(y))-c(F_X(x),F_Y(y))}{c(F_{XT}(x),F_{YT}(y))}\right|\right)\int_{-\infty}^{-c}d\sup_{(x,y)}\sqrt{T}\left|F_{XYT}(x,y)-F_{XY}(x,y)\right| \\
&=o_p(1),
\end{aligned}$$

and by $\sup_{(x,y)}|c(F_{XT}(x),F_{YT}(y))-c(F_X(x),F_Y(y))|=o_p(1)$, $\sup_{(x,y)}\left|c_{k,T}(F_{XT}(x),F_{YT}(y))-c_{k,T}(F_X(x),F_Y(y))\right|$ $C\left(\sup|F_{XT}(x)-F_X(x)|+\sup|F_{XT}(y)-F_X(y)|\right)=o_p(1)$, and the uniform consistency of $c_{k,T}(u_1,u_2)$ to $c(u_1,u_2)$ over $(u_1,u_2)\in(0,1)^2$ [see e.g. Proposition 3 of Bouezmarni, Rombouts and Taamouti

(2010)], and the boundedness of $\int_{-\infty}^{-c} d\sup_{(x,y)}\sqrt{T}\left|F_{XYT}(x,y)-F_{XY}(x,y)\right|$. Hence the Result (i). ■

Proof of Result (ii): Note that by a second order Taylor expansion of the function $g(u)=\log u$ around $u^{*}=1$, we obtain

$$
\begin{aligned}
A_{2T} =& \int\left[\frac{c_{k,T}\left(F_{XT}(x),F_{YT}(y)\right)-c\left(F_X(x),F_Y(y)\right)}{c\left(F_X(x),F_Y(y)\right)}\right]\mathscr{I}\left(x\leq -c, y\leq -c\right)dF_{XY}(x,y)\\
&-\frac{1}{2}\int\left[\frac{c_{k,T}\left(F_{XT}(x),F_{YT}(y)\right)-c\left(F_X(x),F_Y(y)\right)}{c\left(F_X(x),F_Y(y)\right)}\right]^2\mathscr{I}\left(x\leq -c, y\leq -c\right)dF_{XY}(x,y)\\
&+R_t\\
:=& A_{21T}-\frac{1}{2}A_{22T}+R_t\ .
\end{aligned}
$$

where R_t are the remaining terms such that $\sqrt{T}R_T=o_p(1)$.

In the following we show that: **(1)** $\sqrt{T}A_{21T}=T^{-1/2}\sum_{t=1}^{T}\varepsilon_{2t}^{-}+o_p(1)$ and **(2)** $\sqrt{T}A_{22T}=o_p(1)$.

(1) For the first term A_{21T}, noticing that $dF_{XY}(x,y)=c(F_X(x),F_Y(y))dF_X(x)\,dF_Y(y)$ due to the copula representation of the joint cumulative distribution function $F_{XY}(x,y)$, we get

$$
\begin{aligned}
A_{21T} =& \int\left(c_{k,T}(F_{XT}(x),F_{YT}(y))-c(F_X(x),F_Y(y))\right)\mathscr{I}\left(x\leq -c, y\leq -c\right)dF_X(x)\,dF_Y(y)\\
=& \int\left(c_{k,T}(F_{X}(x),F_{Y}(y))-c(F_X(x),F_Y(y))\right)\mathscr{I}\left(x\leq -c, y\leq -c\right)dF_X(x)\,dF_Y(y)\\
&+\int\left(c_{k,T}(F_{XT}(x),F_{YT}(y))-c_{k,T}(F_X(x),F_Y(y))\right)\mathscr{I}\left(x\leq -c, y\leq -c\right)dF_X(x)\,dF_Y(y)\\
:=& A_{211T}+A_{212T}.
\end{aligned}
$$

We shall show that: **(a)** $\sqrt{T}A_{211T}=T^{-1/2}\sum_{t=1}^{T}\varepsilon_{2t}^{-}+o_p(1)$ and **(b)** $\sqrt{T}A_{212T}=o_p(1)$.

(a) By the definition of Bernstein copula density estimator $c_{k,T}(\cdot)$, we have

$$
\begin{aligned}
&\int c_{k,T}\left(F_X(x),F_Y(y)\right)\mathscr{I}\left(x\le -c, y\le -c\right)dF_X(x)\,dF_Y(y)\\
=&\sum_{v_1=0}^{k}\sum_{v_2=0}^{k}C_T\left(\frac{v_1}{k},\frac{v_2}{k}\right)\int P'_{v_1,k}(F_X(x))\mathscr{I}\left(x\le -c\right)dF_X(x)\int P'_{v_2,k}(F_Y(y))\mathscr{I}\left(y\le -c\right)dF_Y(y)\\
=&\sum_{v_1=0}^{k}\sum_{v_2=0}^{k}C_T\left(\frac{v_1}{k},\frac{v_2}{k}\right)\int_0^{F_X(-c)}dP_{v_1,k}(u_1)\int_0^{F_Y(-c)}dP_{v_2,k}(u_2)\\
=&\sum_{v_1=0}^{k}\sum_{v_2=0}^{k}C_T\left(\frac{v_1}{k},\frac{v_2}{k}\right)\big(P_{v_1,k}(F_X(-c))P_{v_2,k}(F_Y(-c))-P_{v_1,k}(F_X(-c))\mathscr{I}(v_2=0)\\
&-P_{v_2,k}(F_Y(-c))\mathscr{I}(v_1=0)+\mathscr{I}(v_1=0)\mathscr{I}(v_2=0)\big)\\
=&\sum_{v_1=0}^{k}\sum_{v_2=0}^{k}C_T\left(\frac{v_1}{k},\frac{v_2}{k}\right)P_{v_1,k}(F_X(-c))P_{v_2,k}(F_Y(-c))\\
=&C_{k,T}\left(F_X(-c),F_Y(-c)\right),
\end{aligned}
$$

where the second equality follows from the change of variables $F_X(x)=u_1$ and $F_Y(y)=u_2$ and the integration by parts, the third equality follows from the fact that $P_{v_1,k}(u_1)\big|_{u_1=0}^{F_X(-c)}=P_{v_1,k}(F_X(-c))-\mathscr{I}(v_1=0)$ and $P_{v_2,k}(u_2)\big|_{u_2=0}^{F_Y(-c)}=P_{v_2,k}(F_Y(-c))-\mathscr{I}(v_2=0)$, the fourth equality from the fact that $C_T(u_1,0)=C_T(0,u_2)=0$, and the last equality from the definition of Bernstein copula distribution estimator.

Now, observing that $\int c\left(F_X(x),F_Y(y)\right)\mathscr{I}\left(x\le -c,y\le -c\right)dF_X(x)\,dF_Y(y)=C\left(F_X(-c),F_Y(-c)\right)$, we get

$$
\begin{aligned}
A_{211T}=&C_{k,T}\left(F_X(-c),F_Y(-c)\right)-C\left(F_X(-c),F_Y(-c)\right)\\
=&C_T\left(F_X(-c),F_Y(-c)\right)-C\left(F_X(-c),F_Y(-c)\right)+o_p\left(T^{-1/2}\right)\\
=&\frac{1}{T}\sum_{t=1}^{T}\{\mathscr{I}\left(F_X(X_t)\le F_X(-c),F_Y(Y_t)\le F_Y(-c)\right)-C\left(F_X(-c),F_Y(-c)\right)\\
&-C_{u_1}\left(F_X(-c),F_Y(-c)\right)\left[\mathscr{I}\left(F_X(X_t)\le F_X(-c)\right)-F_X(-c)\right]\\
&-C_{u_2}\left(F_X(-c),F_Y(-c)\right)\left[\mathscr{I}\left(F_Y(Y_t)\le F_Y(-c)\right)-F_Y(-c)\right]\}+o_p\left(T^{-1/2}\right)\\
=&\frac{1}{T}\sum_{t=1}^{T}\varepsilon_{2t}^{-}+o_p\left(T^{-1/2}\right),
\end{aligned}
$$

under the assumptions of bounded third order partial derivatives for $C(u_1,u_2)$ on $(u_1,u_2)\in(0,1)^2$ and $T^{1/2}k^{-1}\to 0$ as $T\to\infty$. This ends the proof of $\sqrt{T}A_{211T}=T^{-1/2}\sum_{t=1}^{T}\varepsilon_{2t}^{-}+o_p(1)$.

Notice that the last two extra terms appearing in ε_{2t}^{-} are due to the fact that the marginal distributions $F_X(x)$ and $F_Y(y)$ are unknown when estimating the copula function. Furthermore, if $T^{1/2}k^{-1} \to \delta$, for $0 < \delta < \infty$, then there will be an additional bias term with its form depending on the second order partial derives of $C(u_1,u_2)$, i.e. $\partial^2 C(u_1,u_2)/\partial u_l^2$ for $l = 1,2$, appearing in the expression of generalized error ε_{2t}^-. To simplify our theoretical analysis, we focus on the case when the bias term is cancelled by proper choice of k.

(b) Again by the definition of Bernstein copula density estimator $c_{k,T}(\cdot)$, we write

$$
\begin{aligned}
&\int c_{k,T}(F_{XT}(x),F_{YT}(y))\mathscr{I}(x \leq -c, y \leq -c)\, dF_X(x)\, dF_Y(y) \\
=&\int \frac{1}{T}\sum_{t=1}^{T} K_k(F_{XT}(x),F_{YT}(y);F_{XT}(X_t),F_{YT}(Y_t))\mathscr{I}(x \leq -c, y \leq -c)\, dF_X(x)\, dF_Y(y) \\
=&\frac{1}{T}\sum_{t=1}^{T} k^2 \sum_{v_1=0}^{k-1}\sum_{v_2=0}^{k-1} \mathscr{I}(\mathbf{V_t} \in A_k(\mathbf{v})) \int_{-\infty}^{-c} P_{v_1,k-1}(F_{XT}(x))\, dF_X(x) \int_{-\infty}^{-c} P_{v_2,k-1}(F_{YT}(y))\, dF_Y(y). \quad \text{(B.2)}
\end{aligned}
$$

Using a first order Taylor expansion of the function $P_{v_1,k-1}(F_{XT}(x))$ around the true CDF $F_X(x)$, i.e.

$$
P_{v_1,k-1}(F_{XT}(x)) = P_{v_1,k-1}(F_X(x)) + (F_{XT}(x) - F_X(x))\, P'_{v_1,k-1}(F_{XT}(x)) + R_T^1
$$

as well as the expansion for $P_{v_2,k-1}(F_{YT}(y))$, i.e.

$$
P_{v_2,k-1}(F_{YT}(y)) = P_{v_2,k-1}(F_Y(y)) + (F_{YT}(y) - F_Y(y))\, P'_{v_2,k-1}(F_{YT}(y)) + R_T^2,
$$

where R_T^1 and R_T^2 are remaining terms regarding the Taylor expansion of $P_{v_1,k-1}(F_{XT}(x))$ and $P_{v_2,k-1}(F_{YT}(y))$ respectively, with $\sqrt{T}R_T^1 = o_p(1)$ and $\sqrt{T}R_T^2 = o_p(1)$.

we can re-express (B.2) as the summation of the following five terms:

$$
\begin{aligned}
&\frac{1}{T}\sum_{t=1}^{T} k^2 \sum_{v_1=0}^{k-1}\sum_{v_2=0}^{k-1} \mathscr{I}(\mathbf{V_t} \in A_k(\mathbf{v})) \int_{-\infty}^{-c} P_{v_1,k-1}(F_X(x))\, dF_X(x) \int_{-\infty}^{-c} P_{v_2,k-1}(F_Y(y))\, dF_Y(y) \\
&+\frac{1}{T}\sum_{t=1}^{T} k^2 \sum_{v_1=0}^{k-1}\sum_{v_2=0}^{k-1} \mathscr{I}(\mathbf{V_t} \in A_k(\mathbf{v})) \int_{-\infty}^{-c} P_{v_1,k-1}(F_X(x))\, dF_X(x) \int_{-\infty}^{-c} P'_{v_2,k-1}(F_Y(y))(F_{YT}(y) - F_Y(y))\, dF_Y(y) \\
&+\frac{1}{T}\sum_{t=1}^{T} k^2 \sum_{v_1=0}^{k-1}\sum_{v_2=0}^{k-1} \mathscr{I}(\mathbf{V_t} \in A_k(\mathbf{v})) \int_{-\infty}^{-c} P_{v_2,k-1}(F_Y(y))\, dF_Y(y) \int_{-\infty}^{-c} P'_{v_1,k-1}(F_X(x))(F_{XT}(x) - F_X(x))\, dF_X(x) \\
&+\frac{1}{T}\sum_{t=1}^{T} k^2 \sum_{v_1=0}^{k-1}\sum_{v_2=0}^{k-1} \mathscr{I}(\mathbf{V_t} \in A_k(\mathbf{v})) \int_{-\infty}^{-c} P'_{v_1,k-1}(F_X(x))(F_{XT}(x) - F_X(x))\, dF_X(x) \\
&\quad \times \int_{\infty}^{-c} P'_{v_2,k-1}(F_Y(y))(F_{YT}(y) - F_Y(y))\, dF_Y(y) + R_T^1 + R_T^2 \\
=&B_{1T} + B_{2T} + B_{3T} + B_{4T} + R_T^1 + R_T^2.
\end{aligned}
$$

It is straightforward to show that

$$
\begin{aligned}
&\int c_{k,T}(F_X(x),F_Y(y))\mathscr{I}\left(x\le -c, y\le -c\right)dF_X(x)\,dF_Y(y)\\
=&\frac{1}{T}\sum_{t=1}^{T}k^2\sum_{v_1=0}^{k-1}\sum_{v_2=0}^{k-1}\mathscr{I}(\mathbf{V_t}\in A_k(\mathbf{v}))\int_{-\infty}^{-c}P_{v_1,k-1}(F_X(x))\,dF_X(x)\int_{-\infty}^{-c}P_{v_2,k-1}(F_Y(y))\,dF_Y(y)=B_{1T},
\end{aligned}
$$

by simply plugging into the definition of $c_{k,T}(\cdot)$. Therefore, we can decompose the term A_{212T} as follows: $A_{212T}=B_{2T}+B_{3T}+B_{4T}+R_T^1$+$R_T^2$. We next show that $\sqrt{T}B_{jT}=o_p(1)$ for $j=2,3,4$.

Since the analysis of B_{3T} is similar to B_{2T}, in the following we only show the asymptotic negligibility of B_{2T}. Note that by applying the integration by parts, the integral $\int_{-\infty}^{-c}P'_{v_2,k-1}(F_Y(y))(F_{YT}(y)-F_Y(y))\,dF_Y(y)$ is equal to

$$
P_{v_2,k-1}(F_Y(-c))\left(F_{YT}(-c)-F_Y(-c)\right)-\int_{-\infty}^{-c}P_{v_2,k-1}(F_Y(y))\,d\left(F_{YT}(y)-F_Y(y)\right).
$$

Plugging it into the expression of B_{2T}, we obtain

$$
\begin{aligned}
B_{2T}=&\frac{1}{T}\sum_{t=1}^{T}k^2\sum_{v_1=0}^{k-1}\sum_{v_2=0}^{k-1}\mathscr{I}(\mathbf{V_t}\in A_k(\mathbf{v}))\int_{-\infty}^{-c}P_{v_1,k-1}(F_X(x))\,dF_X(x)P_{v_2,k-1}(F_Y(-c))(F_{YT}(-c)-F_Y(-c))\\
&-\frac{1}{T}\sum_{t=1}^{T}k^2\sum_{v_1=0}^{k-1}\sum_{v_2=0}^{k-1}\mathscr{I}(\mathbf{V_t}\in A_k(\mathbf{v}))\int_{-\infty}^{-c}P_{v_1,k-1}(F_X(x))\,dF_X(x)\int_{-\infty}^{-c}P_{v_2,k-1}(F_Y(y))\,d(F_{YT}(y)-F_Y(y))\\
&=B_{21T}-B_{22T}.
\end{aligned}
$$

As $\sqrt{T}(F_{YT}(-c)-F_Y(-c))$ converges in distribution, and therefore is bounded as an $O_p(1)$. Now, noticing that $E\left(\mathscr{I}(\mathbf{V_t}\in A_k(\mathbf{v}))\right)=E\left(\mathscr{I}(\tilde{\mathbf{V}}_\mathbf{t}\in A_k(\mathbf{v}))\right)(1+o(1))=k^{-2}(1+o(1))$, in which the pseudo-observations $\mathbf{V_t}=(F_{XT}(X_t),F_{YT}(Y_t))$ are replaced by the uniformized observations $\tilde{\mathbf{V}}_\mathbf{t}=(F_X(X_t),F_Y(Y_t))$ and hence $\tilde{\mathbf{V}}_\mathbf{t}$, for $t=1,\cdots,T$, are independent and uniformly distributed random variables. By the law of large numbers, we obtain:

$$
\begin{aligned}
&\sqrt{T}B_{21T}\\
=&k^2\sum_{v_1=0}^{k-1}\sum_{v_2=0}^{k-1}\left(\frac{1}{T}\sum_{t=1}^{T}\mathscr{I}(\mathbf{V_t}\in A_k(\mathbf{v}))\right)\int_{-\infty}^{-c}P_{v_1,k-1}(F_X(x))\,dF_X(x)P_{v_2,k-1}(F_Y(-c))\sqrt{T}(F_{YT}(-c)-F_Y(-c))\\
=&\int_{-\infty}^{-c}\left(\sum_{v_1=0}^{k-1}P_{v_1,k-1}(F_X(x))\right)dF_X(x)\left(\sum_{v_2=0}^{k-1}P_{v_2,k-1}(F_Y(-c))\right)\sqrt{T}(F_{YT}(-c)-F_Y(-c))[1+o_p(1)]\\
=&\int_{-\infty}^{-c}dF_X(x)\sqrt{T}(F_{YT}(-c)-F_Y(-c))[1+o_p(1)]\\
=&F_X(-c)\sqrt{T}(F_{YT}(-c)-F_Y(-c))[1+o_p(1)],
\end{aligned}
$$

where the third step term follows from the binomial theorem $\sum_{v_1=1}^{k-1}P_{v_1,k-1}(F_X(x))=\sum_{v_2=1}^{k-1}P_{v_2,k-1}(F_Y(y))=1$ for any x and y.

Similarly, we can show that the term $\sqrt{T}B_{22T}$ satisfies:

$$
\begin{aligned}
&\sqrt{T}B_{22T}\\
=&k^2\sum_{v_1=0}^{k-1}\sum_{v_2=0}^{k-1}\left(\frac{1}{T}\sum_{t=1}^{T}\mathscr{I}(\mathbf{V_t}\in A_k(\mathbf{v}))\right)\int_{-\infty}^{-c}P_{v_1,k-1}(F_X(x))\,dF_X(x)\int_{-\infty}^{-c}P_{v_2,k-1}(F_Y(y))\,d\sqrt{T}(F_{YT}(y)-F_Y(y))\\
=&\int_{-\infty}^{-c}\left(\sum_{v_1=0}^{k-1}P_{v_1,k-1}(F_X(x))\right)dF_X(x)\int_{-\infty}^{-c}\left(\sum_{v_2=0}^{k-1}P_{v_2,k-1}(F_Y(y))\right)d\sqrt{T}(F_{YT}(y)-F_Y(y))[1+o_p(1)]\\
=&\int_{-\infty}^{-c}dF_X(x)\int_{-\infty}^{-c}d\sqrt{T}(F_{YT}(y)-F_Y(y))[1+o_p(1)]\\
=&F_X(-c)\sqrt{T}(F_{YT}(-c)-F_Y(-c))[1+o_p(1)].
\end{aligned}
$$

Combining the above results, we have shown that $\sqrt{T}B_{2T}=\sqrt{T}(B_{21T}-B_{22T})=o_p(1)$. By analogy, we can show $\sqrt{T}B_{3T}=o_p(1)$. The term $\sqrt{T}B_{4T}=o_p(1)$ can also be proved using similar arguments. Thus, we have shown that $\sqrt{T}A_{212T}=o_p(1)$.

(2) To prove that $\sqrt{T}A_{22T}=o_p(1)$, notice that:

$$
\begin{aligned}
&\sqrt{T}A_{22T}\\
=&\sqrt{T}\int\left[\frac{c_{k,T}\left(F_{XT}(x),F_{YT}(y)\right)-c\left(F_X(x),F_Y(y)\right)}{c\left(F_X(x),F_Y(y)\right)}\right]^2\mathscr{I}\left(x\le -c,y\le -c\right)dF_{XY}(x,y)\\
\le&\int\frac{\sqrt{T}\left|c_{k,T}\left(F_{XT}(x),F_{YT}(y)\right)-c\left(F_X(x),F_Y(y)\right)\right|}{c\left(F_X(x),F_Y(y)\right)}dF_{XY}(x,y)\times\sup_{(x,y)\in\mathbb{R}^2}\frac{\left|c_{k,T}\left(F_{XT}(x),F_{YT}(y)\right)-c\left(F_X(x),F_Y(y)\right)\right|}{c\left(F_X(x),F_Y(y)\right)}\\
=&O_p(1)\times o_p(1)=o_p(1),
\end{aligned}
$$

where the third step term follows by using the arguments we used to prove the term A_{21T} and Proposition 3 of Bouezmarni, Rombouts and Taamouti (2010); see also the proof of the term A_{1T}. Hence the Result (ii).

Proof of Theorem 2: We have

$$\sqrt{T}\left[\hat{\rho}^{-}(c)-\hat{\rho}^{+}(c)\right]=\sqrt{T}[\hat{\rho}^{-}(c)-\rho^{-}(c)]-\sqrt{T}[\hat{\rho}^{+}(c)-\rho^{+}(c)]+\sqrt{T}[\rho^{-}(c)-\rho^{+}(c)]$$

From theorem 1, we recall that we have:

$$\begin{aligned}
\sqrt{T}\left[\hat{\rho}^{-}(c)-\rho^{-}(c)\right] &= \tfrac{1}{\sqrt{T}}\textstyle\sum_{t=1}^{T}\left[\varepsilon_{1t}^{-}(c)+\varepsilon_{2t}^{-}(c)\right]+o_p(1), \\
\sqrt{T}\left[\hat{\rho}^{+}(c)-\rho^{+}(c)\right] &= \tfrac{1}{\sqrt{T}}\textstyle\sum_{t=1}^{T}\left[\varepsilon_{1t}^{+}(c)+\varepsilon_{2t}^{+}(c)\right]+o_p(1)
\end{aligned}$$

where

$$\varepsilon_{1t}^{-}(c)=\log\left[c\left(F_X(X_t),F_Y(Y_t)\right)\right]\mathscr{I}\left(X_t\leq -c,Y_t\leq -c\right)-\rho^{-}(c),$$

$$\begin{aligned}
\varepsilon_{2t}^{-}(c) = {} & \mathscr{I}\left(X_t\leq -c,Y_t\leq -c\right)-E\left[\mathscr{I}\left(X_t\leq -c,Y_t\leq -c\right)\right] \\
& -C_{u_1}\left(F_X(-c),F_Y(-c)\right)\left[\mathscr{I}\left(X_t\leq -c\right)-F_X(-c)\right] \\
& -C_{u_2}\left(F_X(-c),F_Y(-c)\right)\left[\mathscr{I}\left(Y_t\leq -c\right)-F_Y(-c)\right],
\end{aligned}$$

$$\varepsilon_{1t}^{+}(c)=\log\left[c\left(F_X(X_t),F_Y(Y_t)\right)\right]\mathscr{I}\left(X_t\geq c,Y_t\geq c\right)-\rho^{+}(c),$$

$$\begin{aligned}
\varepsilon_{2t}^{+}(c) = {} & \mathscr{I}\left(X_t\geq c,Y_t\geq c\right)-E\left[\mathscr{I}\left(X_t\geq c,Y_t\geq c\right)\right] \\
& +C_{u_1}\left(1-F_X(c),1-F_Y(c)\right)\left[\mathscr{I}\left(X_t\leq c\right)-F_X(c)\right] \\
& +C_{u_2}\left(1-F_X(c),1-F_Y(c)\right)\left[\mathscr{I}\left(Y_t\leq c\right)-F_Y(c)\right].
\end{aligned}$$

Following Theorem 1 and replacing $\sqrt{T}[\hat{\rho}^{-}(c)-\rho^{-}(c)]$and $\sqrt{T}[\hat{\rho}^{+}(c)-\rho^{+}(c)]$ by their correspondent expressions, it follows that

$$\begin{aligned}
\sqrt{T}\left[\hat{\rho}^{-}(c)-\hat{\rho}^{+}(c)\right] = {} & \frac{1}{\sqrt{T}}\sum_{t=1}^{T}\left\{\log\left[c\left(F_X(X_t),F_Y(Y_t)\right)\right]\left[\mathscr{I}\left(X_t\leq -c,Y_t\leq -c\right)-\mathscr{I}\left(X_t\geq c,Y_t\geq c\right)\right]\right. \\
& \left.+\left[\varepsilon_{2t}^{-}(c)-\varepsilon_{2t}^{+}(c)\right]\right\}+o_p(1) \\
:= {} & \frac{1}{\sqrt{T}}\sum_{t=1}^{T}u_t(c)+o_p(1).
\end{aligned}$$

where,

$$u_t(c)=\left\{\log\left[c\left(F_X(X_t),F_Y(Y_t)\right)\right]\left[\mathscr{I}\left(X_t\leq -c,Y_t\leq -c\right)-\mathscr{I}\left(X_t\geq c,Y_t\geq c\right)\right]+\left[\varepsilon_{2t}^{-}(c)-\varepsilon_{2t}^{+}(c)\right]\right\}$$

Under the null hypothesis, and using the central limit theorem of the U-statistics $u_t(c)$ for dependent data (see Hall (1984)), we have

$$\sqrt{T}\left[\hat{\rho}^{-}(c)-\hat{\rho}^{+}(c)\right] \to_d N\left(0,\sigma^2(c)\right),$$

where the asymptotic variance $\sigma^2(c)$ is given by

$$\sigma^2(c) = E\left(u_1^2(c)\right) + 2\sum_{j=1}^{\infty} E\left[u_1(c)\,u_{1+j}(c)\right].$$

Proof of proposition 2: First, following similar arguments as in Theorem 1, we can show that the test in (4.4) is consistent, i.e.

$$Pr\left(|\hat{t}(c)| > b\right) \to 1$$

Secondly, following similar arguments as in Theorem 2, and under the local alternative $\rho^{-}(c) = \rho^{+}(c) + \frac{\gamma(c)}{\sqrt{T}}$, and by applying the central limit theorem, we have:

$$\sqrt{T}\left[\hat{\rho}^{-}(c)-\hat{\rho}^{+}(c)\right] \to_d N\left(\gamma(c),\sigma^2(c)\right).$$

■