Dependence Structure of Market States

Desislava Chetalova[*,a], Marcel Wollschläger[a], and Rudi Schäfer[a]

[a]*Faculty of Physics, University of Duisburg-Essen,Lotharstraße 1, 47048 Duisburg, Germany*

February 7, 2020

Revised version published in *Journal of Statistical Mechanics: Theory and Experiment, 2015(8):P08012, 2015.*

Abstract

We study the dependence structure of market states by calculating empirical pairwise copulas of daily stock returns. We consider both original returns, which exhibit time-varying trends and volatilities, as well as locally normalized returns, where the nonstationarity has been removed. The empirical pairwise copula for each state is compared with a bivariate K-copula. This copula arises from a recently introduced random matrix model, in which nonstationary correlations between returns are modeled by an ensemble of random matrices. The comparison reveals overall good agreement between empirical and analytical copulas, especially for locally normalized returns. Still, there are some deviations in the tails. Furthermore, we find an asymmetry in the dependence structure of market states. The empirical pairwise copulas exhibit a stronger lower tail dependence, particularly in times of crisis.

Keywords: Copulas; Market states; Nonstationarity; Asymmetry; Multivariate mixture; K-copula

JEL Classification: C13; C46; C55; G12; G10

1 Introduction

The concept of copulas was introduced by Sklar (1959, 1973) to study the linkage between multivariate distribution functions and their univariate marginals. Since then, copulas have gained growing importance as a tool for modeling statistical dependence of random variables in many fields. In finance, the usage of copulas is relatively new but it has already found applications in risk management (Embrechts et al., 2002, 2003; McNeil et al., 2005; Rosenberg and Schuermann, 2006; Kole et al., 2007; Brigo et al., 2010; Meucci, 2011), derivatives pricing (Rosenberg, 2003; Bennett and Kennedy, 2004; Cherubini et al., 2004; van den Goorbergh et al., 2005; Hull and White, 2006; Hofert and Scherer, 2011), and portfolio optimization (Hennessy and Lapan, 2002; Patton, 2004; Di Clemente and Romano, 2004; Boubaker and Sghaier, 2013). For

[*]Corresponding author. Email: desislava.chetalova@uni-due.de.

an overview of the literature on applications of copulas in finance, the reader is referred to Genest et al. (2009); Patton (2012). Copulas allow to separate the dependence structure of random variables from their marginal distributions. This is sometimes useful in statistical applications as the dependence structure and the marginal distributions can be modeled separately and joined together resulting in new multivariate distributions with different behavior. For a discussion on difficulties in the application of copulas, the reader is referred to Mikosch (2006); Genest and Rémillard (2006); Joe (2006).

Recently, we identified market states as clusters of similar correlation matrices and studied their corresponding correlation structures (Chetalova et al., 2015a). The correlation structure, however, does not capture the full statistical dependence between financial return time series. Here, we choose a copula approach to study the dependence structure of market states. To this end, we calculate empirical copulas for many stock return pairs and average over all of them to obtain an empirical pairwise copula for each market state. We stress that the identification of market states relies on the correlation matrices, the copulas are used only for analyzing the states and not defining them. To estimate the empirical copulas we use both original and locally normalized returns. The original return time series exhibit time-varying trends and volatilities (Black, 1976; Christie, 1982; Schwert, 1989). These have to be taken into account when transforming the marginal distributions to uniform ones. To this end, we apply the method of local normalization by Schäfer and Guhr (2010), which leads to stationary time series while preserving the correlations between them. The resulting empirical copulas provide different information. The copulas for the original returns describe the dependence structure on a global scale, i.e., for the full time horizon, whereas the copulas for the locally normalized returns describe the dependence structure on a local scale.

The empirical pairwise copulas for each market state are compared with a bivariate K-copula introduced by Wollschläger and Schäfer (2016), which arises from a random matrix approach suggested by Schmitt et al. (2013); Chetalova et al. (2015b). It models the nonstationarity of true correlations by an ensemble of random matrices. The model yields a multivariate return distribution in terms of a modified Bessel function of the second kind, a so-called K-distribution. Chetalova et al. (2015a) find that the K-distribution provides a good description of the heavy-tailed empirical return distributions for each market state. Here, we aim to arrive at a consistent description within the random matrix model studying the agreement between K-copula and empirical dependence structure for each market state. In addition, our study provides further evidence for asymmetric dependencies between financial returns (Longin and Solnik, 2001; Ang and Chen, 2002; Hong et al., 2007). We find an asymmetry in the tail dependence of empirical pairwise copulas which we study in more detail.

The paper is organized as follows. In Section 2, we review the basic concepts of copulas, stating the main result in the copula theory, Sklar's theorem, which we use to derive a K-copula. In Section 3, we present the data set and recapitulate the identification of market states for the NASDAQ Composite market in the period 1992–2013 as proposed by Chetalova et al. (2015a). In Section 4, we study the empirical copula densities for each market state and compare them with the K-copula. We conclude our findings in Section 5.

2 Copula

We begin with a short introduction to the concept of copulas in Section 2.1. For more details with an emphasis on the statistical and mathematical foundations of copulas, the reader is referred to the textbooks of Joe (1997) and Nelsen (2006). In Section 2.2, we present the K-copula which plays a central role in this study. We restrict ourselves to the bivariate case since we study empirical pairwise copulas.

2.1 Basic Concepts

Consider two random variables X and Y. The joint distribution of X and Y contains all the statistical information about them. It can be expressed either in terms of the joint probability density function (pdf) $f_{X,Y}(x,y)$ or in terms of the joint cumulative distribution function (cdf) $F_{X,Y}(x,y)$, where

$$F_{X,Y}(x,y) = \int_{-\infty}^{x} \mathrm{d}x' \int_{-\infty}^{y} \mathrm{d}y' \; f_{X,Y}(x',y') \; . \tag{1}$$

From the joint pdf $f_{X,Y}(x,y)$, one can extract the individual distributions of X and Y,

$$f_X(x) = \int_{-\infty}^{\infty} \mathrm{d}y \; f_{X,Y}(x,y) \; , \tag{2}$$

and analogously for Y. The densities $f_X(x)$, $f_Y(y)$, called marginal pdfs, and the corresponding marginal cdfs $F_X(x)$, $F_Y(y)$ describe the individual statistical behavior of the random variables.

When dealing with correlated random variables, one is interested in their statistical dependence. The Pearson correlation coefficient is commonly used as a measure of dependence. It is defined as

$$C_{X,Y} = \frac{\mathrm{Cov}(X,Y)}{\sigma_X \sigma_Y} \; , \tag{3}$$

where $\mathrm{Cov}(X,Y)$ is the covariance between both random variables, and σ_X, σ_Y are the respective standard deviations. However, the correlation coefficient only measures the linear dependence between the random variables.

Copulas provide a natural way to study the statistical dependence of random variables. The transformation

$$U_i = F_i(i) \; , \qquad i = X, Y, \tag{4}$$

leads to new random variables with a uniform distribution on the unit interval, called the rank of X and Y, respectively. Their joint distribution is called a copula. It describes the dependence structure of the random variables X and Y separated from their marginal distributions.

A central result in the copula theory is Sklar's theorem, which enables us to separate any multivariate distribution function into two components: the marginal distribution of each random variable and their copula,

$$F_{X,Y}(x,y) = \mathrm{Cop}_{X,Y}(F_X(x), F_Y(y)) \; . \tag{5}$$

If the marginal cdfs are continuous, the copula that satisfies Equation (5) is given by

$$\mathrm{Cop}_{X,Y}(u,v) = F_{X,Y}(F_X^{-1}(u), F_Y^{-1}(v)) \; , \tag{6}$$

where F_X^{-1} and F_Y^{-1} represent the inverse cdfs, the so-called quantile functions. This equation allows to extract the dependence structure directly from the joint distribution function. From the copula (6), one can compute the copula density as

$$\mathrm{cop}_{X,Y}(u,v) = \frac{\partial^2}{\partial u \partial v} \mathrm{Cop}_{X,Y}(u,v) \; . \tag{7}$$

2.2 K-copula

The K-copula was first introduced by Wollschläger and Schäfer (2016), who find that it describes the empirical dependencies of US stocks much better than a Gaussian copula. The K-copula arises from the random matrix model by Schmitt et al. (2013) to model time-varying correlations between financial time series (Fenn et al., 2011; Münnix et al., 2012). Kremer (2020b) develops the R package kcopula and Kremer (2020a) provides Mathematica code for the bivariate K-copula.

Consider a market consisting of K stocks. At each time $t = 1, \ldots, T$, we assume that the return vector $r(t) = (r_1(t), \ldots, r_K(t))$ is drawn from a multivariate normal distribution with a covariance matrix Σ_t,

$$g(r|\Sigma_t) = \frac{1}{\sqrt{\det 2\pi\Sigma_t}} \exp\left(-\frac{1}{2} r^\dagger \Sigma_t^{-1} r\right) \ , \tag{8}$$

where we suppress the argument t of r to simplify notation. We now model the time-dependent covariance matrix Σ_t by a Wishart random matrix $AA^\dagger$, where the $K \times N$ model matrix A is drawn from a Gaussian distribution with pdf,

$$w(A|\Sigma, N) = \sqrt{\frac{N}{2\pi}}^{KN} \frac{1}{\sqrt{\det \Sigma}^N} \exp\left(-\frac{N}{2} \operatorname{tr} A^\dagger \Sigma^{-1} A\right) \ . \tag{9}$$

Here, Σ represents the average covariance matrix estimated over the sample of all $r(t), t = 1, \ldots, T$. Hence, the time-dependent covariance matrices are modeled by an ensemble of Wishart matrices $AA^\dagger$ which fluctuate around the sample average Σ. Averaging the multivariate normal distribution (8) with the random covariance matrix $AA^\dagger$ over the Wishart ensemble leads to a K-distribution for multivariate returns,

$$\begin{aligned} \langle g\rangle(r|\Sigma, N) &= \int \mathrm{d}[A] \ w(A|\Sigma, N) \ g(r|AA^\dagger) \\ &= \frac{1}{(2\pi)^K \Gamma(N/2)\sqrt{\det \Sigma}} \int_0^\infty \mathrm{d}z \ z^{\frac{N}{2}-1} \mathrm{e}^{-z} \sqrt{\frac{\pi N}{z}}^K \exp\left(-\frac{N}{4z} r^\dagger \Sigma^{-1} r\right) \\ &= \frac{\sqrt{2}^{2-N} \sqrt{N}^K}{\Gamma(N/2)\sqrt{\det(2\pi\Sigma)}} \frac{\mathcal{K}_{\frac{K-N}{2}}\left(\sqrt{N r^\dagger \Sigma^{-1} r}\right)}{\sqrt{N r^\dagger \Sigma^{-1} r}^{\frac{K-N}{2}}} \ , \end{aligned} \tag{10}$$

where $\mathcal{K}_\nu$ is the modified Bessel function of the second kind of order $\nu = (K - N)/2$. It depends only on the average covariance matrix $\Sigma = \sigma C \sigma$ estimated over the whole sample, where C is the average correlation matrix and $\sigma = \operatorname{diag}(\sigma_1, \ldots, \sigma_K)$ the diagonal matrix of standard deviations, and a free parameter N, which governs the variance of the Wishart ensemble,

$$\operatorname{var}([AA^\dagger]_{kl}) = \frac{\Sigma_{kl}^2 + \Sigma_{kk}\Sigma_{ll}}{N} \ , \tag{11}$$

where Σ_{kl} is the kl-th element of the average covariance matrix Σ. Thus, it characterizes the strength of fluctuations around the average covariance matrix in the considered sample. The larger N, the smaller the fluctuations around Σ, eventually vanishing in the limit $N \to \infty$.

The K-copula is the dependence structure which arises for the K-distribution (10). For the bivariate case $K = 2$, the pdf of the vector $r = (r_1, r_2)$ reads

$$f_{c,N}(r_1, r_2) = \langle g\rangle(r|\Sigma, N) = \frac{1}{\Gamma(N/2)} \int_0^\infty \mathrm{d}z \frac{z^{N/2-1}\mathrm{e}^{-z}}{\sqrt{1-c^2}} \frac{N}{4\pi z} \exp\left(-\frac{N}{4z} \frac{r_1^2 - 2c r_1 r_2 + r_2^2}{1-c^2}\right). \tag{12}$$

Here, we used the covariance matrix

$$\Sigma = \begin{pmatrix} \sigma_1^2 & \sigma_1\sigma_2 c \\ \sigma_1\sigma_2 c & \sigma_2^2 \end{pmatrix} = \begin{pmatrix} 1 & c \\ c & 1 \end{pmatrix} \ , \tag{13}$$

where c denotes the average correlation coefficient, estimated over the whole sample. We choose the unit standard deviations, $\sigma_1 = \sigma_2 = 1$, since the copula is independent of the marginal distributions. Then, the marginal pdfs are identical, $f_1(r_1) = f_2(r_2)$, where

$$f_1(r_1) = \int_{-\infty}^{\infty} \mathrm{d}r_2 \; f_{c,N}(r_1, r_2) = \frac{1}{\Gamma(N/2)} \int_{0}^{\infty} \mathrm{d}z \; z^{N/2-1} \mathrm{e}^{-z} \sqrt{\frac{N}{4\pi z}} \exp\left(-\frac{N}{4z} r_1^2\right) \; . \tag{14}$$

According to Equation (6), the bivariate K-copula is given by

$$\mathrm{Cop}_{c,N}(u, v) = F_{c,N}(F_1^{-1}(u), F_2^{-1}(v)) \; , \tag{15}$$

where c and N are the parameters of the copula, $F_{c,N}$ denotes the cdf of the bivariate K-distribution (12), and F^{-1} the inverse marginal cdf given by

$$\begin{aligned} F_{c,N}(r_1, r_2) &= \int_{-\infty}^{r_1} \mathrm{d}\xi \int_{-\infty}^{r_2} \mathrm{d}\zeta \; f_{c,N}(\xi, \zeta) \\ &= \int_{-\infty}^{r_1} \mathrm{d}\xi \int_{-\infty}^{r_2} \mathrm{d}\zeta \int_{0}^{\infty} \frac{\mathrm{d}z}{\Gamma(N/2)} \frac{z^{N/2-1} \mathrm{e}^{-z}}{\sqrt{1-c^2}} \frac{N}{4\pi z} \exp\left(-\frac{N}{4z} \frac{\xi^2 - 2c\xi\zeta + \zeta^2}{1-c^2}\right) \\ &= \int_{-\infty}^{r_1} \mathrm{d}\xi \int_{-\infty}^{r_2} \mathrm{d}\zeta \; \frac{N \sqrt{\frac{N(\xi^2 - 2c\xi\zeta + \zeta^2)}{1-c^2}}^{\frac{N-2}{2}}}{\pi \; \Gamma(N/2) \sqrt{2}^N \sqrt{1-c^2}} \; \mathcal{K}_{\frac{2-N}{2}} \left(\sqrt{\frac{N(\xi^2 - 2c\xi\zeta + \zeta^2)}{1-c^2}} \right) , \end{aligned} \tag{16}$$

and

$$\begin{aligned} F_1(r_1) &= \int_{-\infty}^{r_1} \mathrm{d}\xi \; f_1(\xi) = \int_{-\infty}^{r_1} \mathrm{d}\xi \frac{1}{\Gamma(N/2)} \int_{0}^{\infty} \mathrm{d}z \; z^{N/2-1} \mathrm{e}^{-z} \sqrt{\frac{N}{4\pi z}} \exp\left(-\frac{N}{4z} \xi^2\right) \\ &= \int_{-\infty}^{r_1} \mathrm{d}\xi \; \frac{\sqrt{N} \sqrt{N\xi^2}^{\frac{N-1}{2}}}{\sqrt{\pi} \; \Gamma(N/2) \sqrt{2}^{N-1}} \; \mathcal{K}_{\frac{1-N}{2}} \left(\sqrt{N\xi^2} \right) . \end{aligned} \tag{17}$$

The K-copula density can be obtained from the K-copula (15) by differentiation,

$$\mathrm{cop}_{c,N}(u, v) = \frac{\partial^2 \mathrm{Cop}_{c,N}(u, v)}{\partial u \partial v} = \frac{f_{c,N}(F_1^{-1}(u), F_2^{-1}(v))}{f_1(F_1^{-1}(u)) f_2(F_2^{-1}(v))} \; . \tag{18}$$

It depends only on the average correlation c and the free parameter N, which characterizes the fluctuations around Σ. Figure 1 shows the K-copula density for different parameter values. The stronger the average correlation c and the lower the parameter N, the higher is the probability for extreme co-movements.

Furthermore, we note that the K-copula is a symmetric copula. It is based on the elliptical distribution (10) and thus it belongs to the class of elliptical copulas.

3 Identification of market states

We now present the data set and identify market states as clusters of correlation matrices with similar correlation structures as suggested by Chetalova et al. (2015a). These market states will be the object of our empirical study in the next section.

We analyze daily data of $K = 258$ stocks of the NASDAQ Composite index traded in the 22-year period from January 1992 to December 2013, provided by Yahoo (2014), which corresponds to $T = 5542$ trading days. For each stock k, we calculate the return,

$$r_k(t) = \frac{S_k(t + \Delta t) - S_k(t)}{S_k(t)} \; , \qquad k = 1, \ldots, K, \tag{19}$$

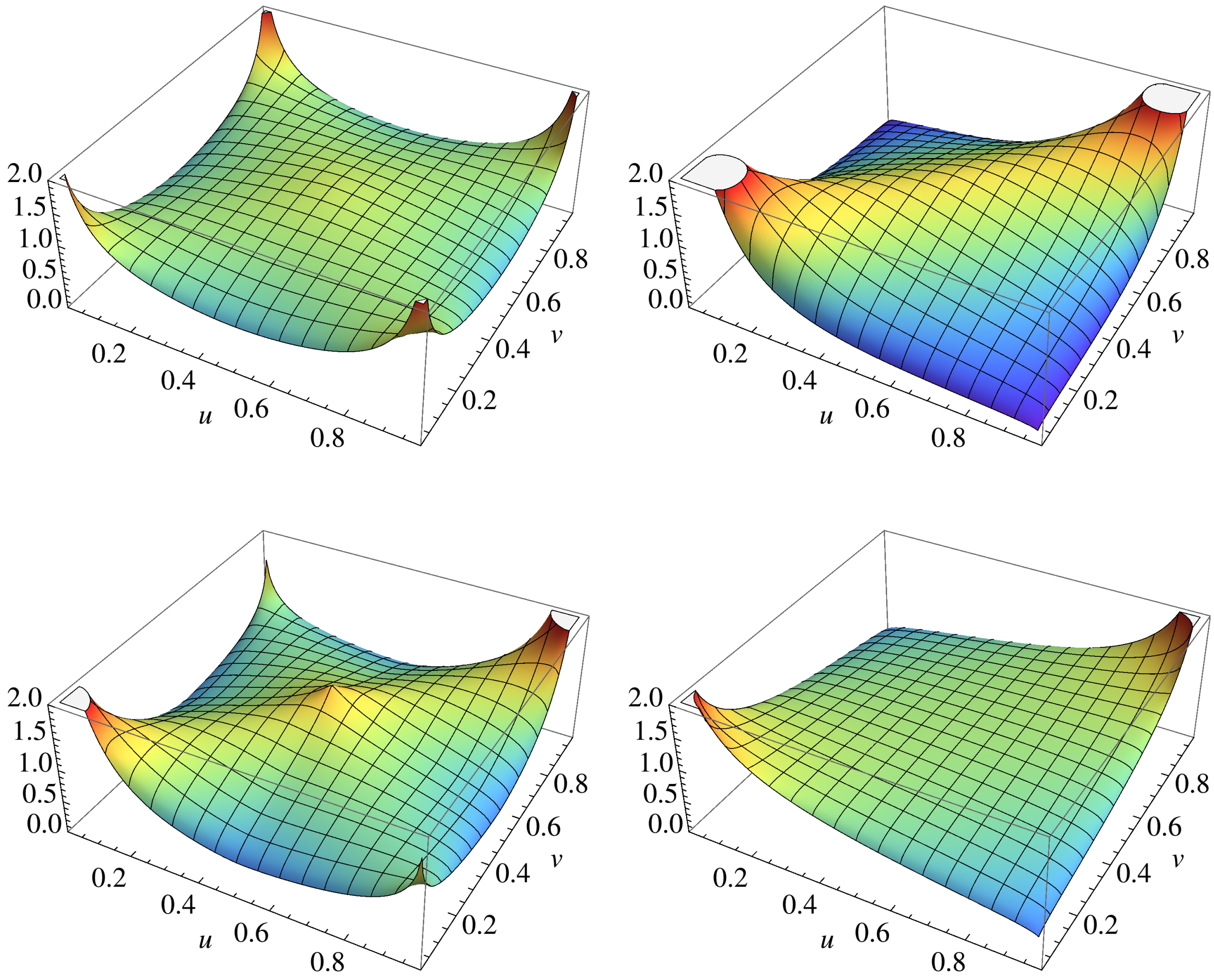

Figure 1: K-copula density $\text{cop}_{c,N}(u, v)$ for different parameter values. Top: $N = 5$ with $c = 0$ (left) and $c = 0.5$ (right), bottom: $c = 0.2$ with $N = 3$ (left) and $N = 30$ (right).

where $S_k(t)$ is the price of the k-th stock at time t and Δt is the return interval which we choose to be one trading day. We refer to the return time series as $r_k = \{r_k(t)\}_{t=1,\ldots,T}$.

Empirical return time series exhibit time-varying drift and volatilities. The correlation coefficient (3) averages over these time-dependent parameters which results in an estimation error for the correlations. In order to eliminate this kind of error, we employ the method of local normalization by Schäfer and Guhr (2010). For each return time series k, we subtract the local mean and divide by the local standard deviation,

$$\hat{r}_k(t) = \frac{r(t) - \langle r(t) \rangle_n}{\sqrt{\langle r^2(t) \rangle_n - \langle r(t) \rangle_n^2}} \,, \tag{20}$$

where $\langle \ldots \rangle_n$ denotes the local average over the n most recent sampling points. For daily data, we use the last $n = 13$ data points (Schäfer and Guhr, 2010). The local normalization removes the local trends and variable volatilities while preserving the correlations between the time series.

Using the locally normalized daily returns, we now obtain a set of 131 correlation matrices calculated on disjoint two-month time intervals of the 22-year observation period. To identify market states, we perform a clustering analysis based on the partitioning around medoids algorithm (Kaufman and Rousseeuw, 1990), where the number of clusters is estimated via the gap statistic (Tibshirani et al., 2001). The clustering analysis separates the set of 131 correlation matrices into six groups based on the similarity of their correlation structures. Each group is associated with a market state. We point out that the identification of market states is performed ex-post, the clustering algorithm has the correlation matrices of all times. Figure 2 shows the time evolution of the six market states. We observe that the market switches back and forth between states. Sometimes it remains in a state for a long time, sometimes it jumps briefly to another state and returns back or evolves further. On longer time scales, the market evolves towards new states, whereas previous states die out.

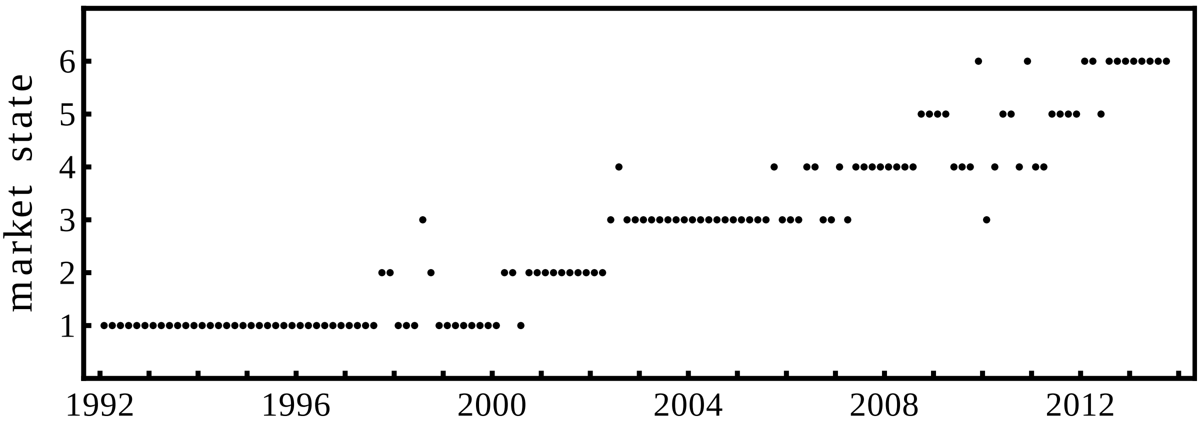

Figure 2: Time evolution of the market in the observation period 1992–2013. Each point represents a correlation matrix measured over a two-month time window.

In the next section, we study the statistical dependence for each market state. To obtain the return time series for each state, we proceed as follows: we take the complete return time series, r_k or $\hat{r}_k$, and divide it into a sequence of disjoint two-month intervals. We merge all intervals belonging to a given state according to the cluster analysis. We note that the return time series for the six market states differ in length.

4 Empirical results

We present the empirical pairwise copula densities for each market states in Section 4.1, and compare them with the K-copula densities in Section 4.2. In Section 4.3 we study the asymmetry of the tail dependence of the empirical copulas in more detail.

4.1 Empirical pairwise copulas for each market state

To calculate the empirical pairwise copula of two return time series r_k and r_l, we first have to transform them into uniformly distributed time series. To achieve this, we employ the empirical distribution function,

$$u_k(t) = F_k(r_k(t)) = \frac{1}{T}\sum_{\tau=1}^{T} \mathbf{1}\{r_k(\tau) \leq r_k(t)\} - \frac{1}{2T} \ , \tag{21}$$

where $\mathbf{1}$ is the indicator function, T denotes the length of the time series, and the factor $1/2$ ensures that the values of the transformed time series u_k lie in the interval $(0, 1)$. The empirical copula density of the time series r_k and r_l can be estimated by the two-dimensional histogram of the transformed data points $(u_k(t), u_l(t))$, $t = 1, \ldots, T$.

An accurate estimation of the copula density requires a large amount of data. Thus, for each state we compute the copula densities of all $K(K-1)/2$ stock pairs and then average over all pairs,

$$\mathrm{cop}^{(i)}(u,v) = \frac{2}{K(K-1)} \sum_{k=1}^{K-1} \sum_{l=k+1}^{K} \mathrm{cop}_{k,l}^{(i)}(u,v) \ , \qquad i = 1, \ldots, 6, \tag{22}$$

where the superscript i denotes the state number. For the bin size of the histograms, we choose $\Delta u = \Delta v = 0.05$. Figure 3 shows the empirical pairwise copula densities for original returns (19) for the six market states. We observe a variation of the dependence structure from state to state, particularly visible in the tails. In state 1, which covers the period from 1992 to roughly 2000, we find a rather flat copula density, indicating low dependence between return pairs. In state 2, we observe deviations from the flat copula density particularly in the tails, which become more and more pronounced in state 3 and 4. State 5, first appearing during the financial crisis in 2008, exhibits the strongest dependence. The dependence decreases again in state 6.

Figure 4 shows the empirical pairwise copula densities for locally normalized returns (20). The dependence structures of the six states are mostly preserved after applying the local normalization. Deviations are observed in the lower-left and the upper-right corners where the copula densities for the original returns exhibit higher peaks. The reason for that is the time-varying volatility. During periods of high volatility, stocks tend to have large returns which contribute to the corners of the copula density. On the other hand, in periods of high volatility the correlations between stocks become stronger. This leads to higher peaks in the corners of the dependence structure for the original returns.

It is important to note that the copulas for original and locally normalized returns contain different statements. The copulas for original returns describe the dependence structure for the full time horizon. On the other hand, the copulas for locally normalized returns provide information about the statistical dependence on a local scale.

Furthermore, we observe that the empirical copula densities are asymmetric with respect to opposite corners. We find stronger dependence in the lower than in the upper tail, that is, the dependence between large negative returns is stronger than the dependence between large positive returns. This asymmetry is an important feature of empirical copula densities and thus we discuss it in more detail in Section 4.3.

4.2 Comparison with the K-copula

We now compare the empirical pairwise copula densities for each market state with the K-copula densities. We consider both original and locally normalized returns. The K-copula density is obtained in the following way: we calculate the K-copula according to equation (15) where the integrals are computed numerically. The K-copula density for each bin of size $\Delta u = \Delta v = 0.05$ is then estimated by

$$\begin{aligned} \mathrm{cop}_{\bar{c},N}(u,v) = \mathrm{Cop}_{\bar{c},N}(u,v) - \mathrm{Cop}_{\bar{c},N}(u,v-\Delta v) - \mathrm{Cop}_{\bar{c},N}(u-\Delta u,v) \\ + \mathrm{Cop}_{\bar{c},N}(u-\Delta u,v-\Delta v) \ . \end{aligned} \tag{23}$$

For comparison, we compute the difference between the empirical and analytical copula density for each state i,

$$\mathrm{cop}^{(i)}(u,v) - \mathrm{cop}_{\bar{c},N}^{(i)}(u,v) \ , \qquad i = 1, \ldots, 6, \tag{24}$$

where $\bar{c}$ is the average correlation coefficient of all $K(K-1)/2$ stock pairs for the considered state. The free parameter N is estimated by a fit which minimizes the mean squared difference between empirical and analytical copula density. The parameter values for each market state are summarized in Table 1. The differences between the empirical copula density and the K-copula density for each state are presented in Figure 5 for original returns and in Figure 6 for locally normalized returns. Overall, we find a good agreement between empirical and analytical

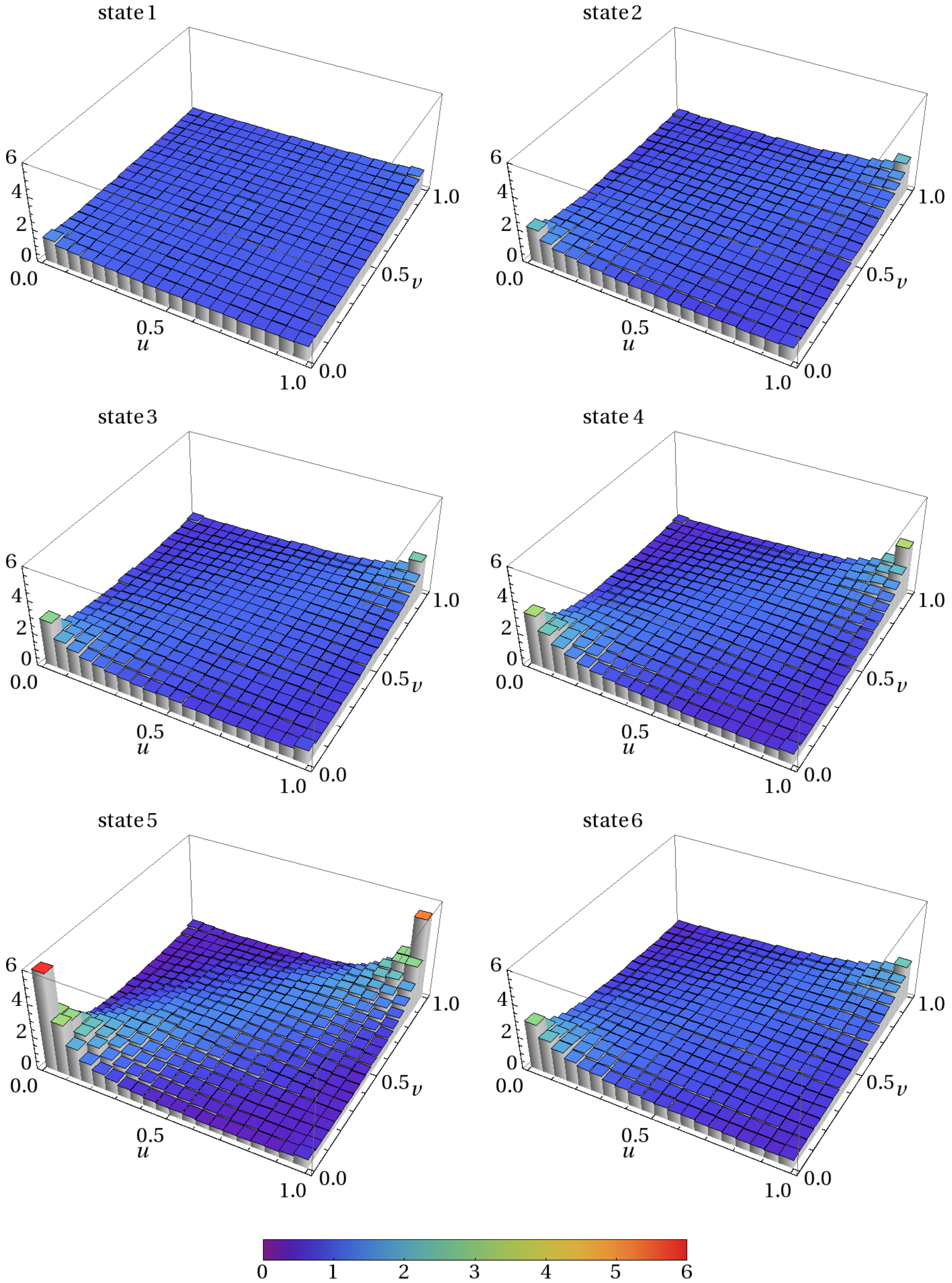

Figure 3: Empirical pairwise copula density $\text{cop}^{(i)}(u, v)$ for market state $i = 1, \ldots, 6$, for original returns.

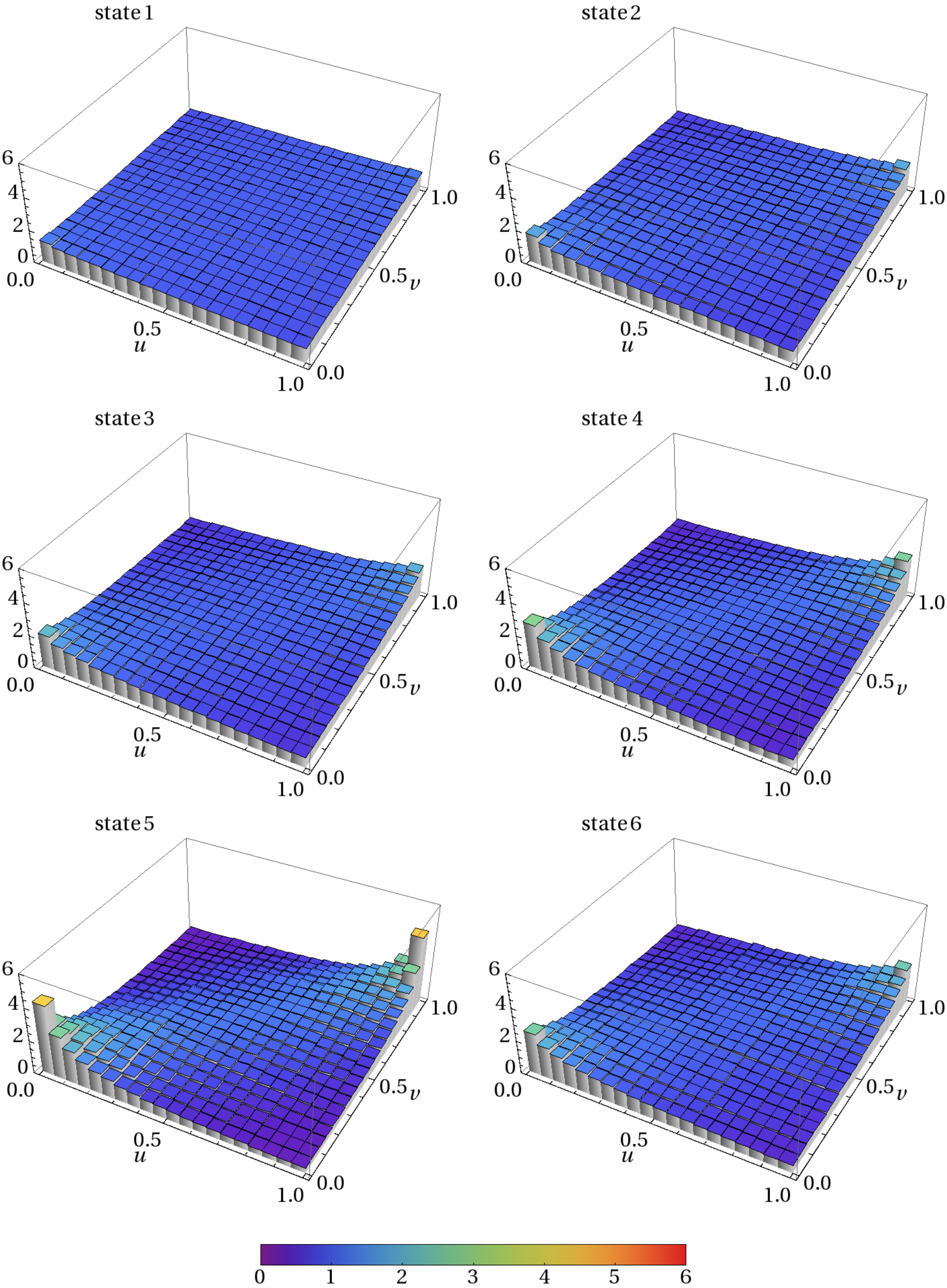

Figure 4: Empirical pairwise copula density $\mathrm{cop}^{(i)}(u, v)$ for market state $i = 1, \ldots, 6$, for locally normalized returns.

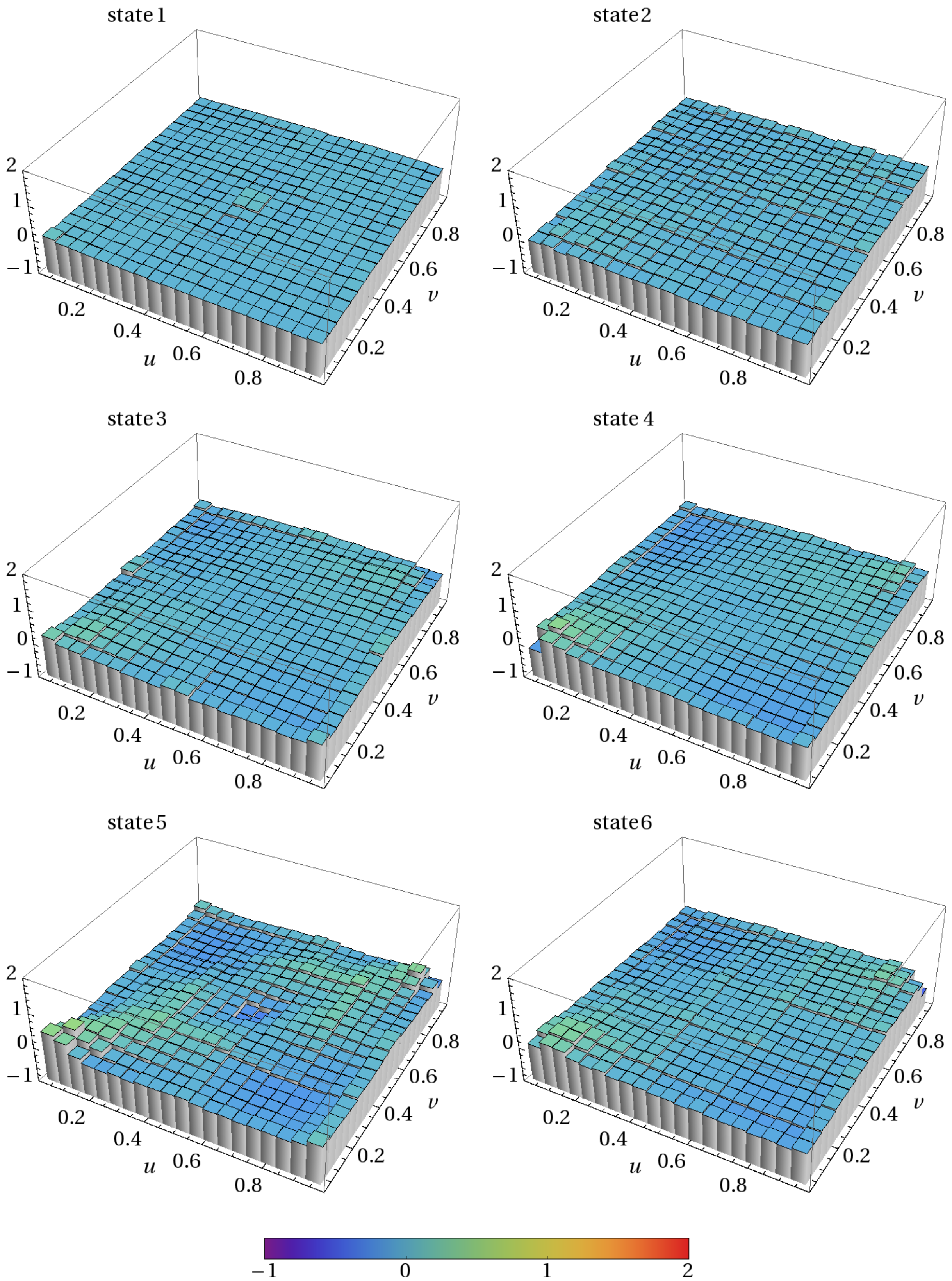

Figure 5: Difference between the empirical copula density and the K-copula density $\mathrm{cop}^{(i)}(u,v) - \mathrm{cop}^{(i)}_{\bar{c},N}(u,v)$ for market state $i = 1, \ldots, 6$, for original returns.

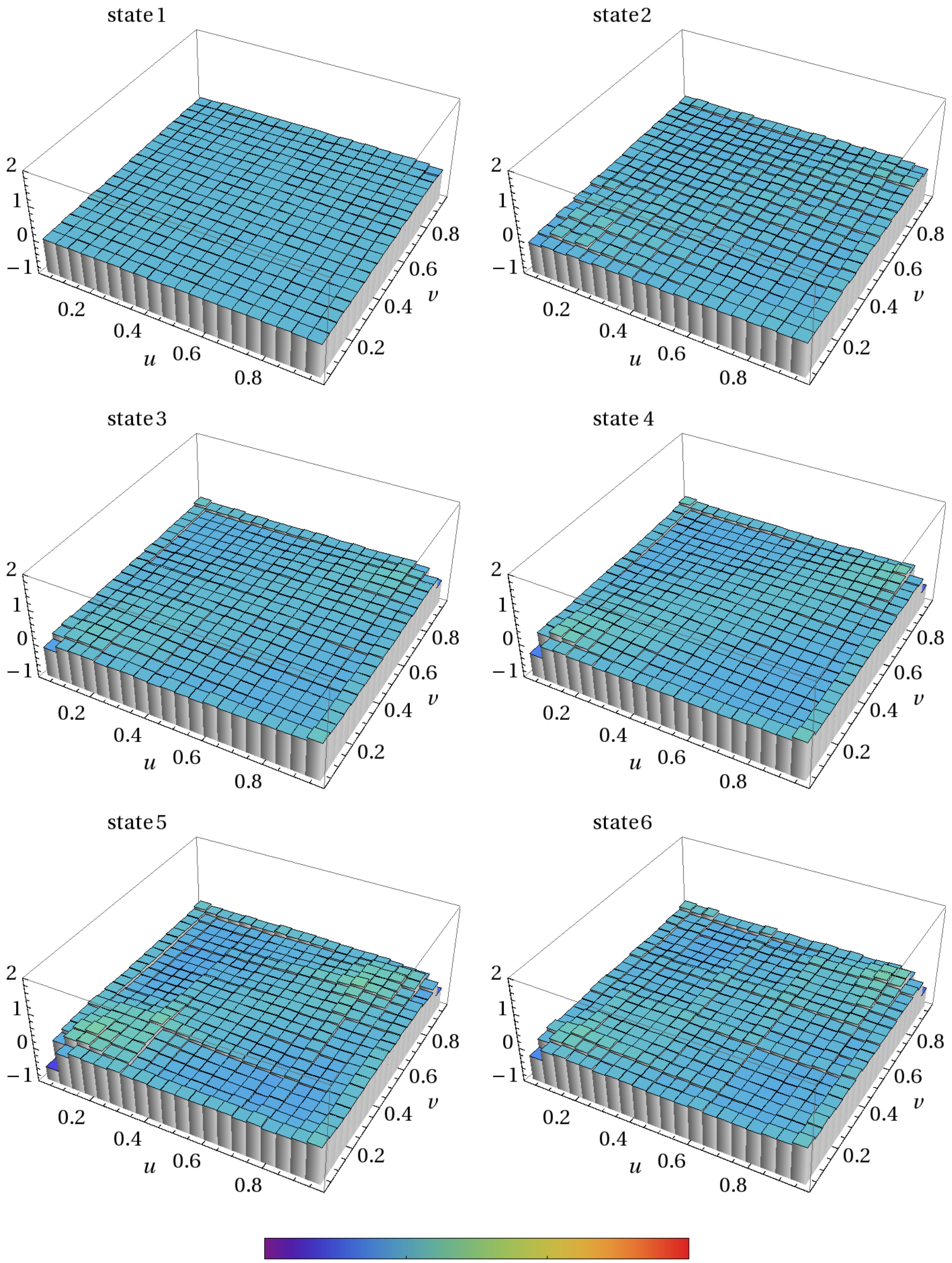

Figure 6: Difference between the empirical copula density and the K-copula density $\mathrm{cop}^{(i)}(u,v) - \mathrm{cop}^{(i)}_{\bar{c},N}(u,v)$ for market state $i = 1, \ldots, 6$, for locally normalized returns.

returns		state 1	state 2	state 3	state 4	state 5	state 6
original	$\bar{c}$	0.046	0.13	0.17	0.25	0.42	0.22
	N	41.7	11.7	8.4	5.6	2.8	10.0
loc. normalized	$\bar{c}$	0.048	0.13	0.19	0.28	0.43	0.25
	N	70.7	28.6	29.8	15.8	7.4	20.4

Table 1: Parameter values of the K-copula density for original and locally normalized returns.

returns	state 1	state 2	state 3	state 4	state 5	state 6
original	0.11	0.41	1.22	2.47	5.53	2.84
loc. normalized	0.047	0.38	0.70	1.43	2.94	1.50

Table 2: Mean squared differences between empirical copula and K-copula densities.

copula densities. The K-copula seems to capture the dependence structure of the first three states very well. Small deviations from the K-copula density are observed for state 4 and 6. Only the dependence structure of state 5 cannot be captured by the K-copula. For locally normalized returns, we find a better agreement which is reflected by the smaller mean squared differences, see Table 2. Deviations are observed mainly in the corners of the copula densities. The empirical copula densities exhibit stronger dependence in the tails, i.e., the probability for extreme co-movements is underestimated by the K-copula.

It is important to note that the K-copula captures the empirical dependence structure quite well with only one free parameter. However, due to its symmetric nature it cannot account for the asymmetry observed in the data. The skewed Student's t-copula is an alternative proposed by Demarta and McNeil (2005) which is able to account for asymmetric dependencies in financial data (Sun et al., 2008; Ammann and Süss, 2009). It captures the empirical dependence structure of original returns better than the K-copula due to the presence of an additional parameter which accounts for the asymmetry (Wollschläger and Schäfer, 2016). Nevertheless, here we confine ourselves to the comparison of the empirical copulas with the K-copula as we aim to arrive at a consistent picture within the random matrix model. We discuss the asymmetry in more detail in the following.

4.3 Asymmetry of the tail dependence

Asymmetric dependencies between returns has been reported by several authors (see e.g., Longin and Solnik, 2001; Ang and Chen, 2002; Hong et al., 2007). Our study provides further evidence revealing a stronger lower tail dependence in the empirical copula densities of market states. We now take a closer look at this asymmetry. To this end, we estimate the tail dependence in the four corners for all $K(K-1)/2$ empirical pairwise copulas,

$$\begin{aligned}
\mathrm{LL}_{k,l}^{(i)} &= \int_{0}^{0.2} \mathrm{d}u \int_{0}^{0.2} \mathrm{d}v\ \mathrm{cop}_{k,l}^{(i)}(u,v)\ , \\
\mathrm{UL}_{k,l}^{(i)} &= \int_{0.8}^{1} \mathrm{d}u \int_{0}^{0.2} \mathrm{d}v\ \mathrm{cop}_{k,l}^{(i)}(u,v)\ , \\
\mathrm{UU}_{k,l}^{(i)} &= \int_{0.8}^{1} \mathrm{d}u \int_{0.8}^{1} \mathrm{d}v\ \mathrm{cop}_{k,l}^{(i)}(u,v)\ , \\
\mathrm{LU}_{k,l}^{(i)} &= \int_{0}^{0.2} \mathrm{d}u \int_{0.8}^{1} \mathrm{d}v\ \mathrm{cop}_{k,l}^{(i)}(u,v)\ ,
\end{aligned} \tag{25}$$

where superscript $i = 1, \ldots, 6$, denotes the state number and the subscripts represent a stock pair k, l. Here, LL and UU refer to the lower-lower and upper-upper corners, respectively, and represent the positive tail dependence, whereas UL and LU refer to the upper-lower and lower-upper corners, respectively, and represent the negative tail dependence. The asymmetry in the tail dependence can be quantified by the differences

$$\begin{aligned}
\alpha_{k,l}^{(i)} &= \mathrm{UU}_{k,l}^{(i)} - \mathrm{LL}_{k,l}^{(i)}\ , \\
\beta_{k,l}^{(i)} &= \mathrm{LU}_{k,l}^{(i)} - \mathrm{UL}_{k,l}^{(i)}\ , \qquad i = 1, \ldots, 6,
\end{aligned} \tag{26}$$

where $\alpha_{k,l}^{(i)}$ captures the asymmetry of the positive tail dependence and $\beta_{k,l}^{(i)}$ the asymmetry of the negative tail dependence for each stock pair k,l. Figure 7 shows the histograms of the asymmetry values for state 5 exemplarily. For the returns, we find a negative offset for the values of $\alpha_{k,l}^{(5)}$, whereas the values of $\beta_{k,l}^{(5)}$ are centered around zero. This indicates, on average, an asymmetry in the positive tail dependence, i.e., simultaneous large negative returns are more likely to occur than simultaneous large positive returns. On the other hand, we do not find such an asymmetry in the negative tail dependence. For the locally normalized returns, we find much weaker asymmetry in the positive tail dependence and no asymmetry for the negative tail dependence. We note that the means of the asymmetry values are the relevant quantities, the asymmetry values for each pair are distributed around the mean due to statistical fluctuations. The standard deviation for $\alpha_{k,l}^{(5)}$ is 0.01 and for $\beta_{k,l}^{(5)}$ 0.007. Indeed, the asymmetry effect is very small. Still, it is clearly visible, see Figure 3.

In the following, we study the asymmetry values for each market state. Figure 8 shows the mean asymmetry values $\bar{\alpha}^{(i)}$ for each market state, obtained by averaging over all $\alpha_{k,l}^{(i)}$ for a given state. On the local scale, the asymmetry in the positive tail dependence is much weaker. Only state 5 still exhibits a certain amount of asymmetry. On the other hand, the asymmetry in the negative tail dependence is negligibly small for both original and locally normalized returns.

5 Conclusion

We study the dependence structure of market states by means of a copula approach. To this end, we estimate the empirical pairwise copulas for each state and compared them with the bivariate K-copula. The bivariate K-copula arises from a random matrix model where the nonstationarity of correlations is taken into account by an ensemble of random matrices. It is a symmetric, elliptical copula which depends on two parameters: the average correlation coefficient, estimated over the considered sample of returns, and a free parameter which characterizes the fluctuations

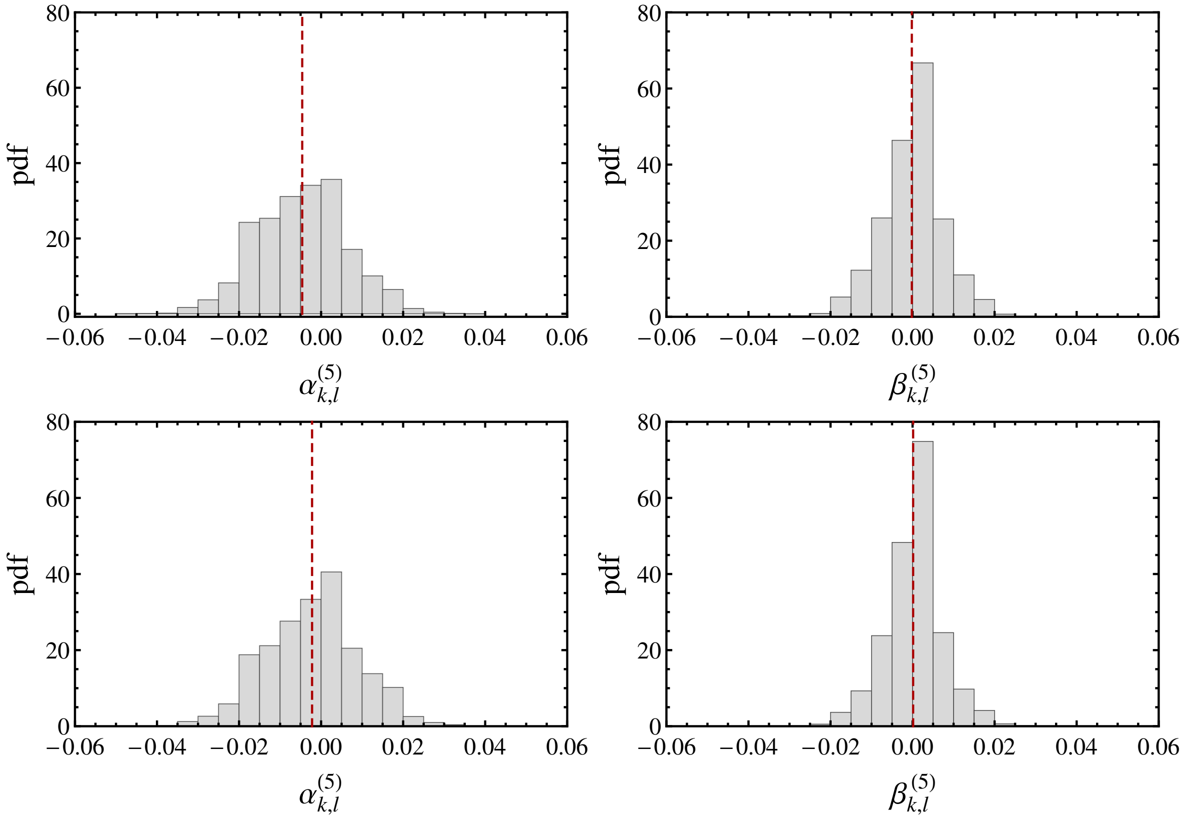

Figure 7: Histograms of the asymmetry values for all stock pairs k, l for state 5. Left: asymmetry for the positive tail dependence $\alpha_{k,l}^{(5)}$, right: asymmetry for the negative tail dependence $\beta_{k,l}^{(5)}$. Top: for original returns, bottom: for locally normalized returns. The dashed red lines represent the corresponding mean values.

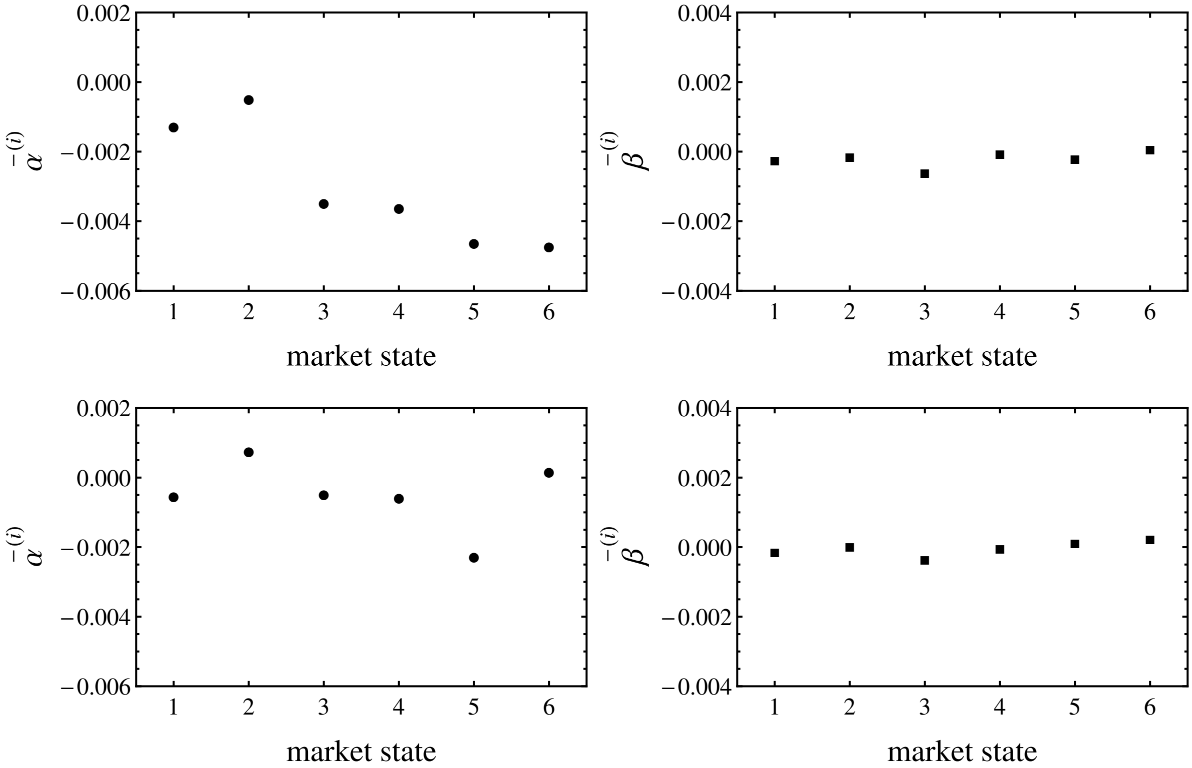

Figure 8: The mean asymmetry values for each state. Left: asymmetry for the positive tail dependence $\bar{\alpha}^{(i)}$, right: asymmetry for the negative tail dependence $\bar{\beta}^{(i)}$. Top: for original returns, bottom: for locally normalized returns.

around the average correlation in this sample. We estimate the empirical pairwise copulas for both original and locally normalized returns. The local normalization removes time-varying trends and volatilities while preserving the dependence structure. The corresponding copula describes the dependence structure on a local scale, whereas the copula of the original returns provides information about the dependence structure on a global scale, i.e., for the full time horizon. Overall, the K-copula captures the empirical dependence structure of market states. We find a good agreement, in particular for the copulas estimated on the local scale. Thus, we obtain a consistent description within the random matrix model: The K-distribution describes the heavy-tailed return distribution while the K-copula captures the corresponding dependence structure. However, we find an asymmetry in the positive tail dependence, i.e., a stronger lower tail dependence, indicating a larger probability for simultaneous extreme negative returns. This asymmetry cannot be captured by our model. It is more pronounced on the global scale. On the local scale, we find a much weaker asymmetry. However, in times of crisis the asymmetry is still clearly present.

References

Ammann, M. and Süss, S. (2009). Asymmetric dependence patterns in financial time series. *European Journal of Finance*, 15:703–719.

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

Bennett, M. N. and Kennedy, J. E. (2004). Quanto pricing with copulas. *Journal of Derivatives*, 12:26–45.

Black, F. (1976). Studies of stock price volatility changes. In *Meetings of the American Statistical Association, Business and Economical Statistics Section*, pages 177–181, Washington DC. American Statistical Association.

Boubaker, H. and Sghaier, N. (2013). Portfolio optimization in the presence of dependent financial returns with long memory: A copula based approach. *Journal of Banking and Finance*, 37:361–377.

Brigo, D., Pallavicini, A., and Torresetti, R. (2010). *Credit Models and the Crisis: A Journey into CDOs, Copulas, Correlations and Dynamic Models.* Wiley Finance, Chichester.

Cherubini, U., Luciano, E., and Vecchiato, W. (2004). *Copula Methods in Finance.* Wiley.

Chetalova, D., Schäfer, R., and Guhr, T. (2015a). Zooming into market states. *J. Stat. Mech.*, P01029.

Chetalova, D., Schmitt, T. A., Schäfer, R., and Guhr, T. (2015b). Portfolio return distributions: Sample statistics with stochastic correlations. *Int. J. Theor. Appl. Finance*, 18:1550012.

Christie, A. A. (1982). The stochastic behavior of common stock variances : Value, leverage and interest rate effects. *J. Finan. Economics*, 10(4):407–432.

Demarta, S. and McNeil, A. (2005). The t copula and related copulas. *International Statistical Review*, 73:111–129.

Di Clemente, A. and Romano, C. (2004). Measuring and optimizing portfolio credit risk: a copula-based approach. *Economic Notes*, 33:325–367.

Embrechts, P., Lindskog, F., and McNeil, A. (2003). Modelling dependence with copulas and application to risk management. In *Handbook of Heavy Tailed Distributions in Finance*, Amsterdam. Elsevier.

Embrechts, P., McNeil, A., and Straumann, D. (2002). Correlation and dependency in risk management: Properties and pitfalls. In *Risk Management: Value at Risk and Beyond*, Cambrige. Cambridge University Press.

Fenn, D. J., Porter, M. A., Williams, S., McDonald, M., Johnson, N. F., and Jones, N. S. (2011). Temporal evolution of financial-market correlations. *Phys. Rev. E*, 84:026109.

Genest, C., Gendron, M., and Bourdeau-Brien, M. (2009). The advent of copulas in finance. *The European Journal of Finance*, 15:609–618.

Genest, C. and Rémillard, B. (2006). Discussion of "Copulas: tales and facts". *Extremes*, 9:27–36.

Hennessy, D. A. and Lapan, H. E. (2002). The use of Archimedean copulas to model portfolio allocations. *Mathematical Finance*, 12:143–154.

Hofert, M. and Scherer, M. (2011). CDO pricing with nested Archimedean copulas. *Quantitative Finance*, 11:775–787.

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

Hull, J. and White, A. (2006). Valuing credit derivatives using an implied copula approach. *The Journal of Derivatives*, 14:8–28.

Joe, H. (1997). *Multivariate Models and Multivariate Dependence Concepts*. Chapman & Hall, London.

Joe, H. (2006). Discussion of "Copulas: tales and facts". *Extremes*, 9:37–41.

Kaufman, L. and Rousseeuw, P. J. (1990). *Finding Groups in Data: An Introduction to Cluster Analysis*. Wiley, New York.

Kole, E., Koedijk, K., and Verbeek, M. (2007). Selecting copulas for risk management. *Journal of Finance and Banking*, 31:2405–2423.

Kremer, M. (2020a). K-copula. `https://github.com/mlkremer/K-copula`. (Accessed on Feb 4, 2020).

Kremer, M. (2020b). kcopula: The bivariate K-copula. `https://github.com/mlkremer/kcopula`. R package version 0.1.0.

Longin, F. M. and Solnik, B. (2001). Extreme correlation of international equity markets. *Journal of Finance*, 56:649–676.

McNeil, A., Frey, R., and Embrechts, P. (2005). *Quantitative Risk Management: Concepts, Techniques and Tools*. Princeton University Press, Princeton, NJ.

Meucci, A. (2011). A new breed of copulas for risk and portfolio management. *Risk*, 24:122–126.

Mikosch, T. (2006). Copulas: tales and facts. *Extremes*, 9:3–20.

Münnix, M. C., Shimada, T., Schäfer, R., Leyvraz, F., Seligman, T. H., Guhr, T., and Stanley, H. E. (2012). Identifying states of a financial market. *Sci. Rep.*, 2:644.

Nelsen, R. B. (2006). *An Introduction to Copulas*. Springer, New York.

Patton, A. J. (2004). On the out-of-sample importance of skewness and asymmetric dependence for asset allocation. *Journal of Financial Economics*, 2:130–168.

Patton, A. J. (2012). Handbook of Economic Forecasting. In *Copula-Based Models for Financial Time Series*, pages 767–785, Berlin. Springer.

Rosenberg, J. V. (2003). Nonparametric pricing of multivariate contingent claims. *Journal of Derivatives*, 10:9–26.

Rosenberg, J. V. and Schuermann, T. (2006). A general approach to integrated risk management with skewed fat-tailed risks. *Journal of Financial Economics*, 79:569–614.

Schäfer, R. and Guhr, T. (2010). Local normalization: Uncovering correlations in non-stationary financial time series. *Physica A*, 389:3856.

Schmitt, T. A., Chetalova, D., Schäfer, R., and Guhr, T. (2013). Non-stationarity in financial time series: Generic features and tail behavior. *Europhys. Lett.*, 103:58003.

Schwert, G. W. (1989). Why does stock market volatility change over time? *J. Finance*, 44:1115–1153.

Sklar, A. (1959). Fonctions de répartition à n dimensions et leurs marges. *Publ. Inst. Statist. Univ. Paris*, 8:229–231.

Sklar, A. (1973). Random variables, joint distribution functions, and copulas. *Kybernetika*, 9:449–460.

Sun, W., Stoyanov, S. V., Rachev, S., and Fabozzi, F. J. (2008). Multivariate skewed students's t copula in the analysis of nonlinear and asymmetric dependence in the German equity market. *Studies in Nonlinear Dynamics and Econometrics*, 12:1572–1572.

Tibshirani, R., Walther, G., and Hastie, T. (2001). Estimating the number of clusters in a data set via the gap statistic. *J. R. Statist. Soc. B*, 63:411–423.

van den Goorbergh, R. W. J., Genest, C., and Werker, B. J. M. (2005). Multivariate option pricing using using dynamic copula models. *Insurance: Mathematics and Economics*, 37:101–114.

Wollschläger, M. and Schäfer, R. (2016). Impact of nonstationarity on estimating and modeling empirical copulas of daily stock returns. *Journal of Risk*, 19(1):1–23.

Yahoo (2014). NASDAQ Composite data. `https://finance.yahoo.com`.