# Dynamic structured copula models

**Wolfgang Karl Härdle, Ostap Okhrin, Yarema Okhrin**

**Angaben zur Veröffentlichung / Publication details:**

Härdle, Wolfgang Karl, Ostap Okhrin, and Yarema Okhrin. 2013. "Dynamic structured copula models." *Statistics & Risk Modeling* 30 (4): 361–88. https://doi.org/10.1524/strm.2013.2004.



Statistics & Risk Modeling 30, 361–388 (2013) / **DOI** 10.1524/strm.2013.2004


# Dynamic structured copula models

**Wolfgang Karl Härdle, Ostap Okhrin*, Yarema Okhrin**

*Received: January 16, 2013; Accepted: August 18, 2013*

**Summary:** There is an increasing demand for models of multivariate time-series with time-varying and non-Gaussian dependencies. The available models suffer from the curse of dimensionality or from restrictive assumptions on the parameters and distributions. A promising class of models is that of hierarchical Archimedean copulae (HAC), which allows for non-exchangeable and non-Gaussian dependency structures with a small number of parameters. In this paper we develop a novel adaptive estimation technique of the parameters and of the structure of HAC for time-series. The approach relies on a local change-point detection procedure and a locally constant HAC approximation. Typical applications are in the financial area but also recently in the spatial analysis of weather parameters. We analyse the time varying dependency structure of stock indices and exchange rates. Both examples reveal periods with constant and turmoil dependencies. The economic significance of the suggested modelling is evaluated using the Value-at-Risk of a portfolio.

## 1 Introduction

The key difference between univariate and multivariate time series analysis is the fact that the future dynamics is affected not only by the univariate past but also by cross-sectional dependencies. These dependencies are not constant and vary in time. Their dynamics, form and strength are important in many applications. Risk diversification, asset allocation, and financial spillovers illustrate this importance. The most straightforward and therefore best established approach to modelling such dependencies is via the correlation (or covariance) matrix. The correlation matrix uniquely characterises the dependency if the data is driven, for example, by a multivariate normal distribution. Similar arguments also hold for arbitrary elliptical distributions (as the multivariate $t$). Due to its simplicity, the covariance matrix has become the standard parametrisation of dependency. In many applications the dependency structure varies over time. Time varying conditional volatilities are modelled using, e.g., GARCH-type processes. For a recent review of multivariate GARCH processes, including Dynamic Conditional Correlation (DCC), Constant Conditional Correlation (CCC), Baba, Engle, Kraft, and Kroner (1990) (BEKK), and others, we refer to Silvennoinen and Teräsvirta (2009). These models still assume that the parameters for the process are constant over an entire estimation period. Such an approach has been

* Corresponding author: Ostap Okhrin
AMS 2010 subject classification: Primary 62H12, Secondary 62H05
Keywords and Phrases: Copula, multivariate distribution, Archimedean copula, adaptive estimation

challenged even in the univariate case, as the growing literature demonstrates, see Lamoureux and Lastrapes (1990). In practice, it is likely that the parameters characterising the dependency change with time in possibly a nonstationary manner.

Another disadvantage of covariance-based dependency modelling is the fact that it fails to capture important types of data features. First, covariances are measures of linear dependence and therefore fail to represent nonlinear relationships. As an alternative approach, one may consider other measures such as Kendall's tau or Spearman's rho, see Joe (1997). However, the extensions of these measures to higher dimensions is problematic, see, e.g., Schmid and Schmidt (2006). Secondly, elliptical distributions postulate symmetric dependency, i.e., the strength of the relationship is the same for high and low values. This is, however, in some applications too restrictive an assumption. Thirdly, the covariance matrix – used as a parameter for a multivariate normal distribution – fails to fit the heavy tails typical of asset returns. An approach which partially solves these problems is based on copulae, proposed by Sklar and reviewed in Joe (1997) and Nelsen (2006). Copulae allow us to model dependency separately from marginal distributions and provide a better fit for heavy tails, asymmetries, etc.

Time-varying copulae were considered recently by Breymann, Dias, and Embrechts (2003), Jondeau and Rockinger (2006), Patton (2004), Rodriguez (2007), and Giacomini, Härdle, and Spokoiny (2009). Patton (2004) considers an asset-allocation problem with a time-varying parameter of bivariate copulae. Rodriguez (2007) studies financial contagion using switching-parameter bivariate copulae. In contrast to those papers, Giacomini, Härdle, and Spokoiny (2009) used a novel method based on local adaptive estimation discussed in Spokoiny (2010). The idea of that approach is to determine a period of homogeneity wherein the parameter of a low-dimensional Archimedean copula (AC) can be approximated by a constant.

The online instantaneous selection of high dimensional dependency structures via multivariate copulae is still an open problem. Here we tackle this problem via multivariate hierarchical Archimedean copulae (HAC). A detailed analysis of this copula class is given in Okhrin, Okhrin, and Schmid (2013a). Unlike simple AC, the HAC is characterised not only by its parameters, but also by its structure. The time-varying dependency therefore affects its structure and parameters simultaneously. The variability of the parameters implies that the dependency becomes stronger or weaker; the variability of the structure implies that there is a change not only in the strength of the dependency, but also in its form. The proposed technique allows us to determine the periods with local constant structure and parameters. It is based on the selection of an appropriate interval out of a set of candidate intervals. This procedure requires the calculation of a sequence of critical values (by simulations) that are used in testing local homogeneity. Local homogeneity is checked via a test against a change point alternative.

To assess the performance of the methodology developed, we perform extensive simulations and empirical studies. Within the simulation study, we show that this novel technique quickly reacts to shifts in the structure and in the parameters. The varying estimation window allows of an increase in the precision of the estimators in stable periods, but simultaneously of reacting quickly to changes if they occur. The detection delay clearly demonstrates the effectiveness of the procedure compared to a rolling

window estimation. In the empirical study, we give one example with changes only in parameters and another example with changes both in the parameters and in the structure.

This paper is structured as follows. In the next section we give a short theoretical background for HAC with estimation and grouping techniques. Section 3 extends the local adaptive estimation procedures to copulae. Sections 4 and 5 deal with applications to simulated and real data.

## 2 Hierarchical Archimedean copulae

The advantage of a copula is that it allows of splitting the multivariate distribution into its margins and a pure dependency component: It captures the dependency between variables, eliminating the impact of the marginal distributions. Formally, copulae were introduced in Sklar (1959). The main result states that if $F$ is an arbitrary $d$-dimensional continuous distribution function of the random variables $X_1, \ldots, X_d$, then the associated copula is unique and defined to be the continuous function $C : [0,1]^d \to [0,1]$ which satisfies the equality

$$C(u_1, \ldots, u_d) = F\{F_1^{-1}(u_1), \ldots, F_d^{-1}(u_d)\}, \quad u_1, \ldots, u_d \in [0,1],$$

where $F_1^{-1}(\cdot), \ldots, F_d^{-1}(\cdot)$ are the quantile functions of the corresponding marginal distributions $F_1(x_1), \ldots, F_d(x_d)$. If $F$ belongs to the class of elliptical distributions, then this results in a so called elliptical copula. Note, however, that this copula cannot be given explicitly, because the distribution function $F$ and the inverse marginal distributions $F_i$ usually have only integral representations. One of the classes that overcomes this drawback of elliptical copulae is the class of Archimedean copulae (AC)

$$C(u_1, \ldots, u_k) = \phi\{\phi^{-1}(u_1) + \cdots + \phi^{-1}(u_d)\}, \quad u_1, \ldots, u_d \in [0,1], \tag{2.1}$$

where $\phi \in \mathfrak{L} = \{\phi : [0;\infty) \to [0,1] \,|\, \phi(0) = 1, \phi(\infty) = 0; (-1)^j \phi^{(j)} \geq 0; j = 1, \ldots, \infty\}$. The function $\phi$ is called the generator of the copula and usually depends on a single parameter $\theta$. For example, the Gumbel generator is given by $\phi = \exp(-x^{1/\theta})$ for $0 \leq x < \infty$, $1 \leq \theta < \infty$. The generator $\phi$ is required to be $d$-monotone, i.e., differentiable up to order $d-2$, with $(-1)^j \phi^{(j)}(x) \geq 0$, $j = 0, \ldots, d-2$ for any $x \in [0,\infty)$, and with $(-1)^{d-2}\phi^{(d-2)}(x)$ non-decreasing and convex on $[0,\infty)$ (e.g., McNeil and Nešlehová (2009)). For a detailed review of the properties of AC can be found in McNeil and Nešlehová (2009) and Joe (1996).

A disadvantage of AC is the fact that the rendered dependency is symmetric with respect to the permutation of variables, i.e., the distribution is exchangeable. Moreover, the multivariate dependency structure is somewhat stiff, since it depends on a single parameter of the generator function $\phi$. The Hierarchical Archimedean Copulae (HAC) overcome this problem by considering the compositions of simple AC. For example, the special case of 4-dimensional HAC fully nested copula can be given by

$$\begin{aligned} C(u_1,u_2,u_3,u_4) &= C_1\{C_2(u_1,u_2,u_3),u_4\} = \phi_1\{\phi_1^{-1} \circ C_2(u_1,u_2,u_3) + \phi_1^{-1}(u_4)\} \\ &= \phi_1\{\phi_1^{-1} \circ \phi_2[\phi_2^{-1}\{C_3(u_1,u_2)\} + \phi_2^{-1}(u_3)] + \phi_1^{-1}(u_4)\}. \end{aligned} \tag{2.2}$$

Composition can be applied recursively using different segmentations of the variables, leading to more complex HACs. For notational convenience let the expression $s = \{(\dots(i_1 \dots i_{j_1})\dots(\dots)\dots)\}$ denote the structure of an HAC, where $i_\ell \in \{1,\dots,d\}$ is a reordering of the indices of the variables. $s_j$ denotes the structure of subcopulae with $s_d = s$. Further, let the $d$-dimensional HAC be denoted by $C(u_1,\dots,u_d;s,\boldsymbol{\theta})$, where $\boldsymbol{\theta}$ the set of copula parameters. For example the fully nested HAC (2.2) can be expressed as

$$\begin{aligned} &C(u_1,\dots,u_d;s=s_d,\boldsymbol{\theta}) \\ &\quad= C\{u_1,\dots,u_d;((s_{d-1})d),(\theta_1,\dots,\theta_{d-1})^\top\} \\ &\quad= \phi_{d-1,\theta_{d-1}}(\phi^{-1}_{d-1,\theta_{d-1}} \circ C\{u_1,\dots,u_{d-1};((s_{d-2})(d-1)),(\theta_1,\dots,\theta_{d-2})^\top\} \\ &\qquad+ \phi^{-1}_{d-1,\theta_{d-1}}(u_d)), \end{aligned}$$

where $s = \{(\dots(12)3)\dots)d)\}$. In Figure 2.1 we present a fully nested HAC with structure $s = (((12)3)4)$ and a partially nested HAC with $s = ((12)(34))$ in dimension $d = 4$.

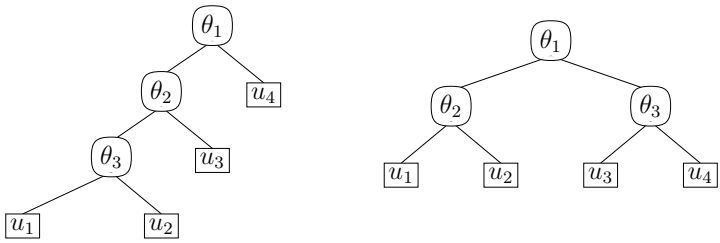


**Figure 2.1** Fully and partially nested copulae of dimension $d = 4$ with structures $s = (((12)3)4)$ on the left and $s = ((12)(34))$ on the right.

HAC are thoroughly analysed in Joe (1997), Whelan (2004), Savu and Trede (2010), Embrechts, Lindskog, and McNeil (2003), Hofert (2012), Hering, Hofert, Mai, and Scherer (2010), etc.

Note that the generators $\phi_i$ within an HAC can come either from a single generator family or from different generator families. If they belong to the same family, then the complete monotonicity of $\phi_i \circ \phi_{i+1}^{-1}$, needed for HAC to be a proper copula, imposes some constraints on the parameters $\theta_1,\dots,\theta_{d-1}$. Theorem 4.4 of McNeil (2008) provides sufficient conditions on the generator functions to guarantee that $C$ is a copula. If $\phi_i \in \mathfrak{L}$ for $i = 1,\dots,d-1$ and $\phi_i \circ \phi_{i+1}^{-1}$ has a completely monotone derivative for $i = 1,\dots,d-2$, then $C$ is a copula. For the majority of generators, a feasible HAC requires decreasing parameters from the lowest to the highest level. However, in the case of different families within a single HAC, the condition of complete monotonicity is more difficult to check.

In general the structure $s$ of the HAC can be arbitrary. On the one hand this makes it a very flexible and at the same time parsimonious distribution model. If we use the same single-parameter generator function on each level, but with a different value of $\theta$, we may specify the whole distribution with $d-1$ parameters. From this point of view, the HAC approach can be seen as a alternative to covariance driven models. On the

other hand, for each HAC not only are the parameters unknown, but also the structure has to be determined. One possible procedure is to enumerate and estimate all possible HACs. Using a suitable goodness-of-fit test we then determine the optimal structure. This approach is however unrealistic in practice even in moderate dimensions. Okhrin, Okhrin, and Schmid (2013a) suggest a computationally efficient procedure, which allows to estimate an HAC recursively.

We restrict the discussion to binary copulae, i.e., at each level of the hierarchy only two variables are joined together. Joining more than two variables dramatically increases the number of formal candidate distributions and the needed computational power. At the lowest level, we fit a bivariate copula to every couple of the variables. The estimation procedure is discussed below. We select the couple of variables with the strongest fit and denote the set of indices of the variables by $I_1$ and the respective estimator of the parameter at the first level by $\hat{\theta}_1 = \hat{\theta}(I_1)$. The selected couple is joined together to define the pseudo-variables $C\{\phi_1, \hat{\theta}_1\}(u_i)_{\{i \in I_1\}}$, i.e we evaluate the copula defined via the generator $\phi_1$ with the parameter $\hat{\theta}$ for the observations of variables from $I_1$. At the next level we proceed in the same way by considering the remaining variables and the new pseudo-variable. This procedure allows us to estimate the structure of the copula.

More formally, the pair $I_j$ to be grouped at the level $j$ is determined from

$$I_j = \underset{I_{k-j+1,i}, |I_{k-j+1,i}|=2}{\operatorname{argmax}} \theta(I_{k-j+1,i}),$$

where $I_{k-j+1,i}$'s are all possible sets of pairs from $k-j+1$ variables at level $j$ and $\theta(I_{k-j+1,i})$ is the corresponding copula parameter. This approach always leads to a feasible copula function with $k-1$ parameters.

If the true copula is not binary, this procedure leads to a potentially mis-specified model. Despite the difference in the structures, the difference in the distribution functions in general is minor. To motivate this point we consider the following binary HAC

$$C_1(\phi, (\theta_1, \theta_2); ((12)3))(u_1, u_2, u_2).$$

If the parameters are close, i.e., for some fixed small $\varepsilon$ : $|\theta_1 - \theta_2| \leq \varepsilon$, then the dependency structure imposed by $C_1$ is very close to the dependency structure imposed by $C_2(\phi, \theta_1; (123))(u_1, u_2, u_2)$. This property is referred to as associativity of AC, see Theorem 4.1.5 of Nelsen (2006). It says that if the parameters are equal then the copula structures (12)3 and (123) lead to the same copula.

Multi-stage maximum-likelihood estimation is a convenient tool for recursive estimation. At the first stage, we estimate the marginal distributions either parametrically or nonparametrically. At the next stage, we estimate the parameter of the copula at the first level replacing marginal distribution functions by their estimates. At further stages, the next level copula parameter is estimated assuming that the margins as well as the copula parameters at lower levels are known.

Let $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_p)^\top$ be the parameters of the copulae starting with the lowest up to the highest level and $\boldsymbol{\alpha} = (\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_d)^\top$ be the vector of parameters of the marginal distributions. The multistage ML estimator $\hat{\boldsymbol{\eta}}$ of $\boldsymbol{\eta} = (\boldsymbol{\alpha}^\top, \boldsymbol{\theta}^\top)^\top$ solves the system

$$\left( \frac{\partial \mathcal{L}_1}{\partial \boldsymbol{\alpha}_1^\top}, \ldots, \frac{\partial \mathcal{L}_d}{\partial \boldsymbol{\alpha}_d^\top}, \frac{\partial \mathcal{L}_{d+1}}{\partial \theta_1}, \ldots, \frac{\partial \mathcal{L}_{d+p}}{\partial \theta_p} \right)^\top = \mathbf{0}, \tag{2.3}$$

$$\begin{aligned}
\text{where} \quad \mathcal{L}_j &= \sum_{i=1}^{n} l_j(\mathbf{X}_i), \text{ for } j = 1, \ldots, d+p, \\
l_j(\mathbf{X}_i) &= \log f_j(x_{ij}, \boldsymbol{\alpha}_j), \text{ for } j = 1, \ldots, d, i = 1, \ldots, n, \\
l_{j+d}(\mathbf{X}_i) &= \log \left( c\big[\{\hat{F}_m(x_{im}, \boldsymbol{\alpha}_m)\}_{m \in s_j}; s_j, \{\theta_\ell\}_{\ell=1,\ldots,p}\big] \prod_{m \in s_j} \hat{f}_m(x_{im}, \boldsymbol{\alpha}_m) \right) \\
&\quad \text{for } j = 1, \ldots, p, \; i = 1, \ldots, n.
\end{aligned}$$

where $c$ is the copula density (see Joe (1997), Savu and Trede (2010)), $\hat{F}_i(\cdot)$ is an estimator of the marginal cdf $F_i$ and $\boldsymbol{X}_j = (x_{1j}, \ldots, x_{nj})^\top$ is the vector of observations of the variable $X_j$. If we estimate the margins parametrically, then $\hat{F}_i(\cdot) = F_i(\cdot, \hat{\boldsymbol{\alpha}}_i)$. The marginal density $\hat{f}(\cdot)$ is estimated accordingly. For the nonparametric estimation of the margins the ML procedure will have the form:

$$\left( \frac{\partial \mathcal{L}_1}{\partial \boldsymbol{\alpha}_1^\top}, \ldots, \frac{\partial \mathcal{L}_p}{\partial \boldsymbol{\alpha}_p^\top} \right)^\top = \mathbf{0}, \tag{2.4}$$

$$\begin{aligned}
\text{where} \quad \mathcal{L}_j &= \sum_{i=1}^{n} l_j(\mathbf{X}_i), \text{ for } j = 1, \ldots, p, \\
l_j(\mathbf{X}_i) &= \log c\big[\{\hat{F}_m(x_{im})\}_{m \in s_j}; s_j, \{\theta_\ell\}_{\ell=1,\ldots,p}\big] \\
&\quad \text{for } j = 1, \ldots, p, \; i = 1, \ldots, n.
\end{aligned}$$

Chen and Fan (2006) and Okhrin, Okhrin, and Schmid (2013a) provide the asymptotic behaviour of the estimates.

# 3 Inhomogeneous dependence

Numerous models have been proposed for time-varying correlation structure, with the multivariate GARCH model being among the most popular ones. In these models, the correlations are defined as functions of (lagged) explanatory variables which may influence the variation in the current dependency structure (measured via the correlation). This implies that the conditional correlation changes at each moment of time, but the parameters of the conditioning functions are assumed to be constant. This approach suffers from two important drawbacks. First, the estimation of this type of model is tedious because of the large number of parameters to be estimated. Second, there is evidence that the parameters do change with time, see, e.g., Lamoureux and Lastrapes (1990). Neglecting this fact may lead to inconsistent estimators.

In order to cope with a time varying dependency structure, we propose a parsimonious alternative that is based on a locally constant HAC approximation. With a once-and-for-all calculated set of critical values, we determine the periods of homogeneity instantaneously at each time point. The corresponding theory and applications may be found in Čížek, Härdle, and Spokoiny (2009), Mercurio and Spokoiny (2004), and Chen, Härdle, and Jeong (2008). This method can be applied to virtually any dependency model. When applied to HACs, it allows us to control not only for periods with constant parameters, but also for periods with constant structure. Moreover, this method is applicable not only to abrupt changes in the dependency, but also to smooth transitions in the model parameters.

Let $s_t$ and $\boldsymbol{\theta}_t$ denote the time varying but unknown copula structure and parameters. The idea is to select for each time $t_0$ an interval $I$ in which $\boldsymbol{\theta}_t$ and $s_t$ can be well approximated by constant $\boldsymbol{\theta}^*$ and $s^*$. The discrepancy between two copulae $C(\cdot;s,\boldsymbol{\theta})$ and $C(\cdot;s',\boldsymbol{\theta}')$ is measured by the Kullback–Leibler (KL) divergence $\mathcal{K}$:

$$\mathcal{K}\{C(\cdot;s',\boldsymbol{\theta}'),C(\cdot;s,\boldsymbol{\theta})\} = \mathsf{E}_{s',\boldsymbol{\theta}'} \log \frac{c(\cdot;s',\boldsymbol{\theta}')}{c(\cdot;s,\boldsymbol{\theta})}$$

where $c$ is the copula density, see Nelsen (2006). The aim is to select $I$ as close as possible to the so-called 'oracle' choice interval. Define the 'oracle' choice $I_{k^*}$ of interval as the largest interval $I = [t_0 - m_{k^*}; t_0]$ for which the small modelling bias condition (SMB) is fulfilled:

$$\triangle_I(s,\boldsymbol{\theta}) = \sum_{t\in I} \mathcal{K}\{C(\cdot;s_t,\boldsymbol{\theta}_t),C(\cdot;s,\boldsymbol{\theta})\} \leq \triangle, \text{ for some } \triangle \geq 0,\ s,\boldsymbol{\theta}. \tag{3.1}$$

The unknown local constant copula parameter (at $t_0$) can be best estimated on the largest interval $\operatorname{argmax}_I \triangle_I(s,\boldsymbol{\theta}) = [t_0 - m_{k^*}; t_0]$ fulfilling (3.1). This means that on the interval $I$, the model $s_t$, $\boldsymbol{\theta}_t$ can be well approximated by a constant $s$, $\boldsymbol{\theta}$. The methods of estimation of the HAC discussed in Okhrin, Okhrin, and Schmid (2013a) maximise the ML with respect to the structure $s$ and parameters $\boldsymbol{\theta}$, which leads to the best parametric fit to the underlying model on $I$, defined by $\tilde{s}_I$, $\tilde{\boldsymbol{\theta}}_I$. Recall that the Kullback–Leibler divergence plays a particular role in the analysis and estimation of mis-specified models, see White (1982). In the case of minimising $\triangle_I(s,\boldsymbol{\theta})$ with respect to the length of the interval $I$, we minimise the loss caused by the ignorance of the time variation in the copula.

Note that the true time-varying parameters $\boldsymbol{\theta}_t$ and $s_t$ are unknown. Therefore also the 'oracle' choice $m_{k^*}$ is unknown. In a data-driven algorithm based on the Local Change Point (LCP) detection procedure, see Spokoiny (2010), we sequentially test whether $\boldsymbol{\theta}_t = \boldsymbol{\theta}^*$ and the structure of the HAC $s_t = s^*$ is constant within some interval $I$. Here the aim of the LCP technique is not to detect a change point, but rather to conveniently determine the period of constant dependency. Alternative techniques can be found, for example in Čížek, Härdle, and Spokoiny (2009).

The risk arising in the estimation of locally constant copulae under the SMB is bounded. Let $\mathcal{L}(s,\boldsymbol{\theta})$ denote the log-likelihood function based on the HAC with the

parameters $s$ and $\boldsymbol{\theta}$. Following Čížek, Härdle, and Spokoiny (2009), let $\tilde{\boldsymbol{\theta}}_I$ and $\tilde{s}_I$ be any estimators on an interval $I$. If the SMB holds for some $I$, $s$, and $\boldsymbol{\theta}$, then

$$\mathsf{E}_{s_t,\boldsymbol{\theta}_t} \log\left\{1+\frac{|\mathcal{L}(\tilde{s}_I,\tilde{\boldsymbol{\theta}}_I)-\mathcal{L}(s,\boldsymbol{\theta})|^r}{\Re_r(s,\boldsymbol{\theta})}\right\} \le 1+\triangle, \tag{3.2}$$

where $\Re_r(s,\boldsymbol{\theta})$ is an upper bound satisfying

$$\mathsf{E}_{s^*,\boldsymbol{\theta}^*}|\mathcal{L}(\tilde{s}_I,\tilde{\boldsymbol{\theta}}_I)-\mathcal{L}(s^*,\boldsymbol{\theta}^*)|^r \le \Re_r(s^*,\boldsymbol{\theta}^*)$$

which is called a 'propagation condition'. We set $r=0.5$ since this choice is also proposed in the literature. The bound given in (3.2) tells us that the risk in an estimated local constant model (under SMB) differs from the risk in the true constant model by $\triangle$.

The LCP is based on sequentially testing the hypotheses of homogeneity on intervals $I_k$. We select $I_k$ with $k=-1,0,1,\ldots$ as a sequence of intervals $I_k \subset I_{k+1}$, starting with $k=1$. If there are no change points in $\mathcal{T}_k \subset I_k \setminus I_{k-1}$, then we accept $I_k$ as an interval with constant copula parameter and structure. At the next step we take $\mathcal{T}_{k+1}$ and test it for homogeneity. We repeat these steps until rejection or the largest possible interval $I_K$ is accepted, leading to an interval $I_{\hat{k}}$.

Two sources of error occur in practical applications. Let $I_{k^*}$ denote the oracle choice. This implies that for $I_k$ $(k<k^*)$ the SMB holds. The first type of error ('false alarm') occurs if $\hat{k}<k^*$. In this case the estimation is based on a shorter data period and therefore implies higher variability. Let $\hat{s}_k$ and $\hat{\boldsymbol{\theta}}_k$ be the respective estimators and $\tilde{s}_k$ and $\tilde{\boldsymbol{\theta}}_k$ denote the corresponding fixed-sample estimators on $I_k$. Under the SMB condition on $I_{k^*}$ and assuming that $\max_{k\le k^*}\mathsf{E}_{s,\boldsymbol{\theta}}|\mathcal{L}(\tilde{s}_k,\tilde{\boldsymbol{\theta}}_k)-\mathcal{L}(s,\boldsymbol{\theta})|^r \le \Re_r(s,\boldsymbol{\theta})$, we obtain by Theorem 4.2 of Čížek, Härdle, and Spokoiny (2009)

$$\begin{aligned}
&\mathsf{E}_{s_t,\boldsymbol{\theta}_t} \log\left\{1+\frac{|\mathcal{L}(\tilde{s}_{\hat{k}},\tilde{\boldsymbol{\theta}}_{\hat{k}})-\mathcal{L}(s,\boldsymbol{\theta})|^r}{\Re_r(s,\boldsymbol{\theta})}\right\} \le 1+\triangle, \qquad (3.3)\\
&\mathsf{E}_{s_t,\boldsymbol{\theta}_t} \log\left\{1+\frac{|\mathcal{L}(\tilde{s}_{\hat{k}},\tilde{\boldsymbol{\theta}}_{\hat{k}})-\mathcal{L}(\hat{s}_{\hat{k}},\hat{\boldsymbol{\theta}}_{\hat{k}})|^r}{\Re_r(s,\boldsymbol{\theta})}\right\} \le \rho+\triangle.
\end{aligned}$$

The inequalities (3.3) say that if we observe a false alarm at the step $\hat{k}<k^*$, then the estimation risk measured by $|\mathcal{L}(\tilde{s}_{\hat{k}},\tilde{\boldsymbol{\theta}}_{\hat{k}})-\mathcal{L}(\hat{s}_{\hat{k}},\hat{\boldsymbol{\theta}}_{\hat{k}})|^r$ is of the same order as the risk of a pure parametric estimation with fixed interval given by $I_{\hat{k}}$.

The second type of error arises if $\hat{k}>k^*$. Outside the oracle interval we are exploiting data which does not support the SMB condition. This implies that the bounds in (3.3) increase. Theorem 4.3 of Čížek, Härdle, and Spokoiny (2009) provides general bounds for the adaptive estimator, showing that

$$\mathsf{E}_{s_t,\boldsymbol{\theta}_t} \log\left\{1+\frac{|\mathcal{L}(\tilde{s}_{k^*},\tilde{\boldsymbol{\theta}}_{k^*})-\mathcal{L}(\hat{s}_{\hat{k}},\hat{\boldsymbol{\theta}}_{\hat{k}})|^r}{\Re_r(s,\boldsymbol{\theta})}\right\} \le 1+\triangle+\log\left\{1+\frac{\mathfrak{z}^r_{k^*}}{\Re_r(s,\boldsymbol{\theta})}\right\}, \tag{3.4}$$

where $\mathfrak{z}_{\hat{k}}$ are the critical values of the test for homogeneity and are defined below. This statement implies that the copula based on $\hat{s}_{\hat{k}}$ and $\hat{\boldsymbol{\theta}}_{\hat{k}}$ belongs with high probability to the confidence interval of the oracle copula with $\tilde{s}_{k^*}$ and $\tilde{\boldsymbol{\theta}}_{k^*}$.

## 3.1 Local test of homogeneity

A local homogeneity test can now be performed. Let us fix some $t_0$ and let $I = [t_0 - m, t_0]$ be an interval candidate and $\mathcal{T}_I$ be a set of interval points within $I$. We estimate the copula parameter $\theta$ and the structure $s$ from observations in $I$, assuming the homogeneous model within $I$, i.e., using the notation from the previous section, $\hat{\boldsymbol{\theta}}_{t_0} = \tilde{\boldsymbol{\theta}}_I$ and $\hat{s}_{t_0} = \tilde{s}_I$. The null hypothesis $H_0$ means that $\forall \tau \in \mathcal{T}_I : \boldsymbol{\theta}_\tau = \boldsymbol{\theta}$, $s_\tau = s$, i.e., the observations in $I$ follow the model with the dependence parameter $\boldsymbol{\theta}$ and the structure $s$. The alternative (change point) hypothesis $H_1$ claims that $\exists \tau \in \mathcal{T}_I : \boldsymbol{\theta}_t = \boldsymbol{\theta}_1$ and $s_t = s_1$ for $t \in J = [\tau, t_0]$ and $\boldsymbol{\theta}_t = \boldsymbol{\theta}_2 \neq \boldsymbol{\theta}_1$ or $s_t = s_2 \neq s_1$ for $t \in J^c = [t_0 - m, \tau)$, i.e., either the parameter $\boldsymbol{\theta}$ or the whole structure $s$ changed spontaneously at some intermediate point $\tau$ of the interval $I$. In other words,

$$
\begin{aligned}
H_0 &: \forall \tau \in \mathcal{T}_I,\ \boldsymbol{\theta}_t = \boldsymbol{\theta},\ s_t = s,\ \forall t \in I = J \cup J^c = [\tau, t_0] \cup [t_0 - m, \tau) \\
H_1 &: \exists \tau \in \mathcal{T}_I,\ \boldsymbol{\theta}_t = \boldsymbol{\theta}_1,\ s_t = s_1;\ \forall t \in J = [\tau, t_0], \\
&\quad \text{and } \boldsymbol{\theta}_t = \boldsymbol{\theta}_2 \neq \boldsymbol{\theta}_1 \text{ or } s_t = s_2 \neq s_1, \forall t \in J^c = [t_0 - m, \tau).
\end{aligned}
$$

If $\mathcal{L}_I(s, \boldsymbol{\theta})$ and $\mathcal{L}_J(s_1, \boldsymbol{\theta}_1) + \mathcal{L}_{J^c}(s_2, \boldsymbol{\theta}_2)$ are the log-likelihood functions corresponding to $H_0$ and $H_1$, respectively, the likelihood ratio test for the single change point with known fixed location $\tau$ is given by

$$
T_{I,\tau} = \max_{s_1, \boldsymbol{\theta}_1, s_2, \boldsymbol{\theta}_2} \{\mathcal{L}_J(s_1, \boldsymbol{\theta}_1) + \mathcal{L}_{J^c}(s_2, \boldsymbol{\theta}_2)\} - \max_{s, \boldsymbol{\theta}} \mathcal{L}_I(s, \boldsymbol{\theta}).
$$

Since the point $\tau$ is unknown, one defines the test statistics:

$$
T_I = \max_{\tau \in \mathcal{T}_I} T_{I,\tau}.
$$

It tests the homogeneity hypothesis in $I$ against a change point alternative with unknown location $\tau$ (in the set $\mathcal{T}_I$). The decision rule of the test requires comparing $T_I$ with the critical value $\mathfrak{z}_I$. The critical value depends on the interval $I$, the dimension, and the parameter of the copula. We reject the hypothesis of homogeneity if $T_I > \mathfrak{z}_I$. To run the test, several parameters have to be specified. This includes the choice of the interval candidates $(I_k)$ and internal points $\mathcal{T}_k = \mathcal{T}_{I_k}$ for each of this intervals and the choice of the critical values $\mathfrak{z}_k = \mathfrak{z}_{I_k}$. One possible example of the implementation is based on the choice of the interval candidates $(I_k)$ in the form of a geometric grid. If the length of the interval $I_1$ is fixed at $m_1$, then we define $m_0 = a_1 m_1$ and $m_{-1} = a_2 m_1$ for $a_1 > a_2 \in (0, 1)$ and $m_k = [m_1 c^{k-1}]$ for $k = 1, 2, \ldots, K$ and $c > 1$, where $[x]$ means the integer part of $x$. Further we set $I_k = [t_0 - m_k, t_0]$ and $\mathcal{T}_k = [t_0 - m_{k-1}, t_0 - m_{k-2}]$ for $k = 1, 2, \ldots, K$, see Figure 3.1.

For each particular copula model and for each sequence of intervals the critical values $\mathfrak{z}_k$ are determined from simulations. Under the null hypothesis of homogeneous dependence, the adaptive estimator should coincide with the largest allowed interval $I_K$. However, if the estimated interval is $I_{\hat{k}}$ with $\hat{k} < K$, then the test rejects a correct null hypothesis. The critical values are therefore determined not from the classical level condition, but relying on the precision of the parameter estimators. If $\hat{k}$ is small, the

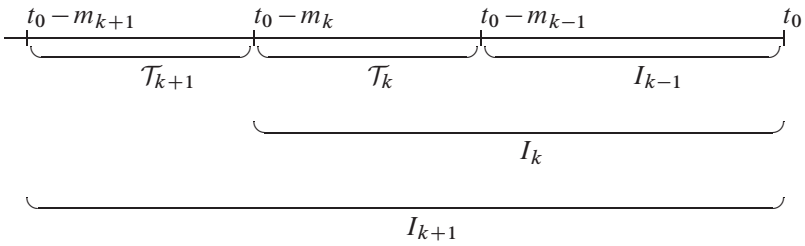


**Figure 3.1** Interval selection.

volatility of the parameter estimator is high. This implies that false decisions with small $\hat{k}$ more severely deteriorate the test of homogeneity. To overcome this problem, we ensure that

$$\mathsf{E}_{s^*,\theta^*}|\mathcal{L}(\tilde{s}_k,\tilde{\boldsymbol{\theta}}_k)-\mathcal{L}(\hat{s}_k,\hat{\boldsymbol{\theta}}_k)|^r \le \rho_k \Re_r(s^*,\boldsymbol{\theta}^*),$$

where $\rho_k = \rho k/K \le \rho$ and $\Re_r(s^*,\boldsymbol{\theta}^*) = \max_k |\mathcal{L}(\tilde{s}_k,\tilde{\boldsymbol{\theta}}_k)-\mathcal{L}(s^*,\boldsymbol{\theta}^*)|^r$. The parameter $\rho$ plays the role of the level of significance and influences the sensitivity of the procedure to inhomogeneity.

In this paper we used the sequential choice of critical values $\mathfrak{z}_k$ discussed in Spokoiny (2009). In the situation after $k$ steps of the algorithm, we distinguish between two cases. In the first, the change point is detected at some step $\ell \le k$ and in the other case no change point is detected. Following the notation in Spokoiny (2009), let $\mathcal{B}_\ell$ be the event meaning the rejection of the null hypothesis at step $\ell$

$$\mathcal{B}_\ell = \{T_1 \le \mathfrak{z}_1, \ldots, T_{\ell-1} \le \mathfrak{z}_{\ell-1}, T_\ell > \mathfrak{z}_\ell\},$$

and $(\hat{s}_k,\hat{\boldsymbol{\theta}}_k) = (\tilde{s}_{\ell-1},\tilde{\boldsymbol{\theta}}_{\ell-1})$ on $\mathcal{B}_\ell$ for $\ell = 1,\ldots,k$. By Monte Carlo simulations from some fixed parametric models discussed in Section 4, we found sequentially the minimal value of $\mathfrak{z}_l$ that ensures the following inequality:

$$\max_{k=l,\ldots,K} \mathsf{E}_{s^*,\theta^*}|\mathcal{L}(\tilde{s}_k,\tilde{\boldsymbol{\theta}}_k)-\mathcal{L}(\tilde{s}_{\ell-1},\tilde{\boldsymbol{\theta}}_{\ell-1})|^r \mathbf{I}(\mathcal{B}_\ell) \le \rho \Re_r(s^*,\boldsymbol{\theta}^*)k/(K-1),$$

where $\mathbf{I}$ is the indicator function. For $\ell = 1$ this inequality depends only on $\mathfrak{z}_1$ in $\mathcal{B}_1 = \{T_1 > \mathfrak{z}_1\}$. For every $\ell \ge 2$ we take $\mathfrak{z}_1,\ldots,\mathfrak{z}_{\ell-1}$ fixed from the previous step, which means that $\mathcal{B}_\ell$ is controlled only by $\mathfrak{z}_\ell$. Throughout the study, we fix $r = 0.5$. Large values of $\rho$ lead to smaller critical values, and small $\rho$ to larger.

# 4 Simulation study

How quickly does the LCP react to shifts in the parameters and/or in the structure? We consider a 3-dimensional HAC with Gumbel generators and uniform margins. To simulate from an HAC, we used the algorithm of McNeil (2008). We consider samples of size 400,

where a change in parameters and structure occurs at $t = 200$. The parameter changes are modelled by

$$C_t(u_1,u_2,u_3;s,\boldsymbol{\theta}) = \begin{cases} C\{u_1, C(u_2,u_3;\theta_1 = 3.33);\theta_2 = 1.43\} & \text{for } 1 \le t \le 200, \\ C\{u_1, C(u_2,u_3;\theta_1 = 2.00);\theta_2 = 1.43\} & \text{for } 200 < t \le 400; \end{cases} \tag{4.1}$$

$$C_t(u_1,u_2,u_3;s,\boldsymbol{\theta}) = \begin{cases} C\{u_1, C(u_2,u_3;\theta_1 = 3.33);\theta_2 = 1.43\} & \text{for } 1 \le t \le 200, \\ C\{u_1, C(u_2,u_3;\theta_1 = 3.33);\theta_2 = 2.00\} & \text{for } 200 < t \le 400. \end{cases} \tag{4.2}$$

Via model (4.1), we investigate the sensitivity of a downward jump in $\theta_1$, while (4.2) is designed for study of an upward jump in $\theta_2$. The initial parameters $\theta_1 = 3.33$ and $\theta_2 = 1.43$ correspond to Kendall $\tau$ equal to 0.7 and 0.3, respectively. In (4.1), $\tau_1$ decreases to 0.5 from 0.7, while in (4.2) $\tau_2$ increases to 0.5 from 0.3. Note that in both cases the difference between the parameters becomes smaller.

The change point in the structure is modelled in a similar way:

$$C_t(u_1,u_2,u_3;s,\boldsymbol{\theta}) = \begin{cases} C\{u_1, C(u_2,u_3;\theta_1 = 3.33);\theta_2 = 1.43\} & \text{for } 1 \le t \le 200, \\ C\{C(u_1,u_2;\theta_1 = 3.33),u_3;\theta_2 = 1.43\} & \text{for } 200 < t \le 400. \end{cases} \tag{4.3}$$

Our technique is implemented with a family of interval candidates of a geometric grid form defined by $m_0 = 20, 40$ and $c = 1.25$. The values of $m_0$ and $c$ have turned out to provide stable results, which is confirmed in the literature cited earlier. Note that the simulated critical values are indifferent to the form of the initial structure $s_1 = ((12)3)$ or $s_2 = (1(23))$, but depend on the parameters. Using the fact that for a Gumbel copula the parameter $\theta \in [1;\infty)$ is unbounded from above, we define the grid based on Kendall's $\tau$ by

$$\boldsymbol{\theta} = (\theta_1,\theta_2)^\top = \{\theta(\tau_1),\theta(\tau_2)\}^\top,$$

where

$$\{\tau_1,\tau_2\} \in \{0.1,0.3,0.5,0.7,0.9\}^2,\ \tau_1 \ge \tau_2.$$

This grid in correlation space corresponds to the grid in parameter space given by $\{\theta_1,\theta_2\} \in \{1.11,1.43,2,3.33,10\}^2,\ \theta_1 \ge \theta_2$. Thus, we simulate from copulae $C\{u_1, C(u_2,u_3;\theta_1 = 3.33);\theta_2 = 1.43\}$, $C\{u_1, C(u_2,u_3;\theta_1 = 2.00);\theta_2 = 1.43\}$, etc. The case $\theta_1 = \theta_2$ corresponds to the simple 3-dimensional AC $C(u_1,u_2,u_3,\theta_1)$. To estimate ${}_{\delta k}\ k = 1,\dots,K = 10$, we simulate $N = 10000$ samples of size $n = [m_0 c^K] + 1$ using the same geometric grid of the intervals. Knowing the true pair $(s^*,\boldsymbol{\theta}^*)$ we used for simulations, we calculate $\Re_r(s^*,\boldsymbol{\theta}^*)$ for each sample as the largest absolute deviation (to the power $r$) of the likelihood with ML estimates $(\tilde{s}_k,\tilde{\boldsymbol{\theta}}_k)$ over the given interval $k$ from the likelihood with given true parameters $(s^*,\boldsymbol{\theta}^*)$.

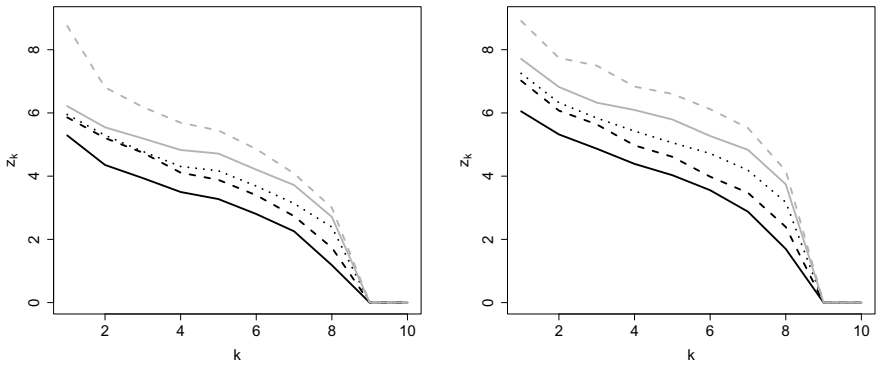


**Figure 4.1** Simulated critical values of the 3-dimensional Gumbel HAC as a function of $k$ with the parameters of the geometric grid set to $m_0 = 20$ (left) and $m_0 = 40$ (right). $\rho$ and $\tau_1$ are fixed and equal to 0.5 and 0.1, respectively, while $\tau_2$ varies: $\tau_2 = 0.1$ (solid black), $\tau_2 = 0.3$ (dashed black), $\tau_2 = 0.5$ (dotted black), $\tau_2 = 0.7$ (solid grey), $\tau_2 = 0.9$ (dashed grey).

Figure 4.1 shows the behaviour of the critical values as a function of $k$ for different values of $\tau_2$.

Usually, it is difficult to work with grids in higher dimensions, as the number of possible models increases exponentially. In this case we propose an on-line calculation of the critical values for every $\{\tilde{\boldsymbol{\theta}}_I, \tilde{s}_I, \}$. The critical values are not calculated in advance, but are computed for each occurring parameter constellation individually. This makes the whole procedure computationally more expensive, but still appropriate in higher dimensions. Furthermore, the procedure is more precise, since we use the exact parameter values, instead of approximations via the grid points. Thus small changes in the parameters should be also identifiable. Nevertheless, our simulation study and empirical analysis are in three dimensions, which allows us to use the grid.

In each change point model we simulate $n = [m_0 c^K] + 400$ observations, where the first $[m_0 c^K]$ values are used as a pre-run for model estimation. For each $t = t_0$ starting from $[m_0 c^K] + 1$ we apply the LCP to the recent observations, i.e., we determine the interval with constant dependency and estimate the corresponding HAC. The results are shown in Figures 4.3, 4.4 and 4.5. $m_0$ is set to 20 in the left column and to 40 in the right column, $\rho$ is set to 0.5. The solid lines show the average values, the dashed line the median values and the grey areas show the interval containing 95 of 100 replications.

The shifts in the first parameter for (4.1) and in the second parameter for (4.2) are plotted in the first rows of Figures 4.3 and 4.4 respectively. Figure 4.5 illustrates the application of LCP to the change-point model (4.3), where in the first row we show the changes in the structure and in the second row the changes in the parameters. For all three types of shift, we observe that the average estimated parameter or structure smoothly moves from the value before the shift to the value after the shift. The delayed reaction naturally depends on $m_0$. Smaller values of $m_0$ allow our procedure to react more quickly to changes. On the other hand the precision of the estimation decreases with decreasing sample size.

The last two rows of all three figures show the dynamics of the average length of the estimated interval of homogeneity and the behaviour of the maximum-likelihood. The estimation is initiated with the shortest available interval of homogeneity. Since the copula is stable and more observations become available, the length of the interval increases to the largest allowed value. After the shift the length of the interval decreases and increases only after the change-point leaves the smallest allowed interval.

| | | $m_0 = 20$ | | | | | $m_0 = 40$ | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| model | r | Q1 | Med. | Mean | Q3 | SD | Q1 | Med. | Mean | Q3 | SD |
| | 0.25 | 2.00 | 8.0 | 11.74 | 19.25 | 12.00 | 8.00 | 21.0 | 22.65 | 31.25 | 18.59 |
| (4.1) | 0.50 | 12.00 | 18.0 | 20.86 | 28.00 | 13.52 | 28.00 | 37.0 | 39.83 | 46.25 | 20.09 |
| | 0.75 | 16.75 | 27.0 | 30.75 | 39.00 | 22.51 | 37.00 | 52.5 | 59.53 | 73.75 | 31.83 |
| | 0.25 | 0.00 | 9.0 | 13.20 | 20.00 | 14.23 | 8.00 | 28.5 | 30.87 | 50.25 | 24.14 |
| (4.2) | 0.50 | 8.75 | 25.0 | 25.02 | 35.00 | 19.41 | 36.75 | 50.0 | 52.58 | 63.25 | 26.47 |
| | 0.75 | 19.00 | 31.0 | 35.19 | 47.00 | 25.77 | 50.00 | 63.5 | 72.71 | 87.00 | 35.18 |
| (4.3) | | 15.00 | 18.0 | 17.78 | 21.00 | 5.23 | 28.00 | 32.0 | 32.45 | 37.00 | 5.90 |

**Table 4.1** Detection delay statistics for LCP, $\rho = 0.5$.

| | | rolling window | | | | |
|---|---|---|---|---|---|---|
| model | r | Q1 | Med. | Mean | Q3 | SD |
| | 0.25 | 10.00 | 26.0 | 26.69 | 40.25 | 19.12 |
| (4.1) | 0.50 | 36.75 | 57.5 | 55.90 | 74.75 | 24.38 |
| | 0.75 | 77.00 | 95.5 | 95.53 | 112.50 | 29.82 |
| | 0.25 | 6.25 | 35.5 | 39.57 | 64.50 | 34.52 |
| (4.2) | 0.50 | 51.00 | 75.5 | 76.64 | 103.00 | 38.16 |
| | 0.75 | 85.25 | 113.0 | 109.70 | 128.20 | 35.90 |
| (4.3) | | 68.00 | 75.5 | 75.53 | 84.25 | 11.40 |

**Table 4.2** Detection delay statistics for a rolling window with length of the estimation period equal to the average length of the intervals of homogeneity.

The detection ability of the proposed procedure is conveniently characterised by the detection delay. Denote by $\gamma_i$ the size of the jump at time $t = 200$, i.e., $\gamma_i = \theta_{it} - \theta_{i,t-1}$ with $i = 1$ for the model (4.1) and $i = 2$ for the model (4.2). The detection delay $\delta$ at rule $\mathrm{r} \in [0, 1]$ is defined by

$$\delta(t, \gamma_i, \mathrm{r}) = \min\{u \geq t : \hat{\theta}_{iu} - \theta_{i,t-1} \geq \mathrm{r}|\gamma_i|\} - t$$

and shows the number of steps needed to detect the fraction r of the jump in the true $\theta$. For the model (4.3) we just look for the time point after $t = 200$ where the structure $s_2 = (1(23))$ is obtained for the first time

$$\delta(t) = \min\{u \geq t : \hat{s}_u \neq s_{t-1}\} - t.$$

Spokoiny (2010) argued that the detection delays are proportional to the probability of an error of type II. Table 4.1 represents the descriptive statistics of the detection delay for

different models (4.1), (4.2) and (4.3), $\mathfrak{r} \in \{0.25, 0.5, 0.75\}$ and $m_0 \in \{20, 40\}$. To detect half of the shift in the parameters, the procedure needs 20 to 25 observations for $m_0 = 20$ and 40 to 50 observations for $m_0 = 40$. The detection ability of the procedure for changes in the structure is similar. We also observe that the mean detection delay is higher for upward jumps than for downward jumps. The mathematical reason for this is explained below. Table 4.2 contains the detection delays for the rolling window estimation. To make the comparison fair, we set the length of the estimation window equal to the average length of the intervals of homogeneity in the LCP procedure. We observe that the flexible choice of the interval of homogeneity leads to substantially shorter detection delays, compared to the rolling estimation.

To get more insight into detection delay we consider the difference

$$\begin{aligned} &\mathcal{K}[C\{s_0; \theta(0.1, 0.2)^\top\}, C\{s_0; \theta(\tau_1, \tau_2)^\top\}] \\ &\quad - \mathcal{K}[C\{s_0; \theta(\tau_1, \tau_2)^\top\}, C\{s_0; \theta(0.1, 0.2)^\top\}], \end{aligned} \tag{4.4}$$

where $\theta(\tau_1, \tau_2)$ denotes the vector of parameters corresponding to the Kendall $\tau$s given by $\tau_1$ and $\tau_2$. The first term in (4.4) denotes the KL divergence between the true copula

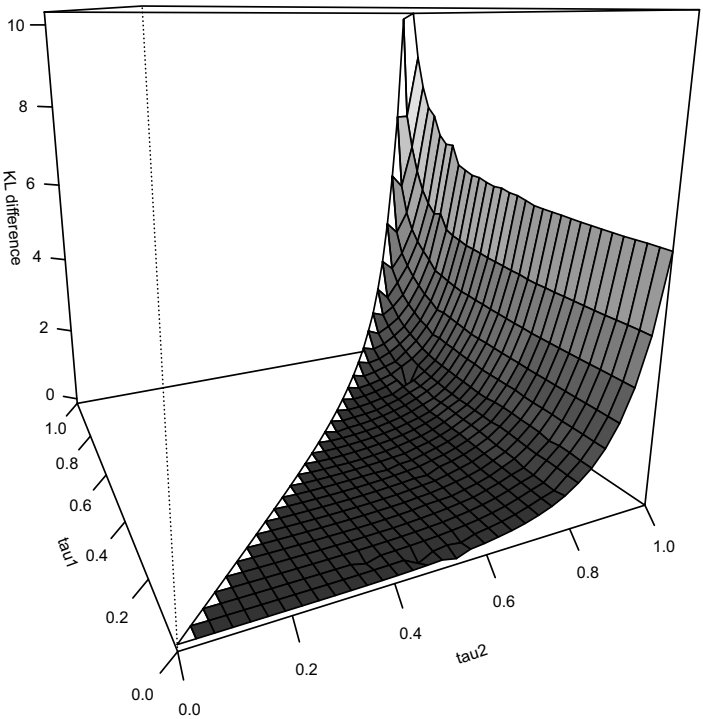


**Figure 4.2** The difference between the KL divergences for mis-specified models in (4.4).

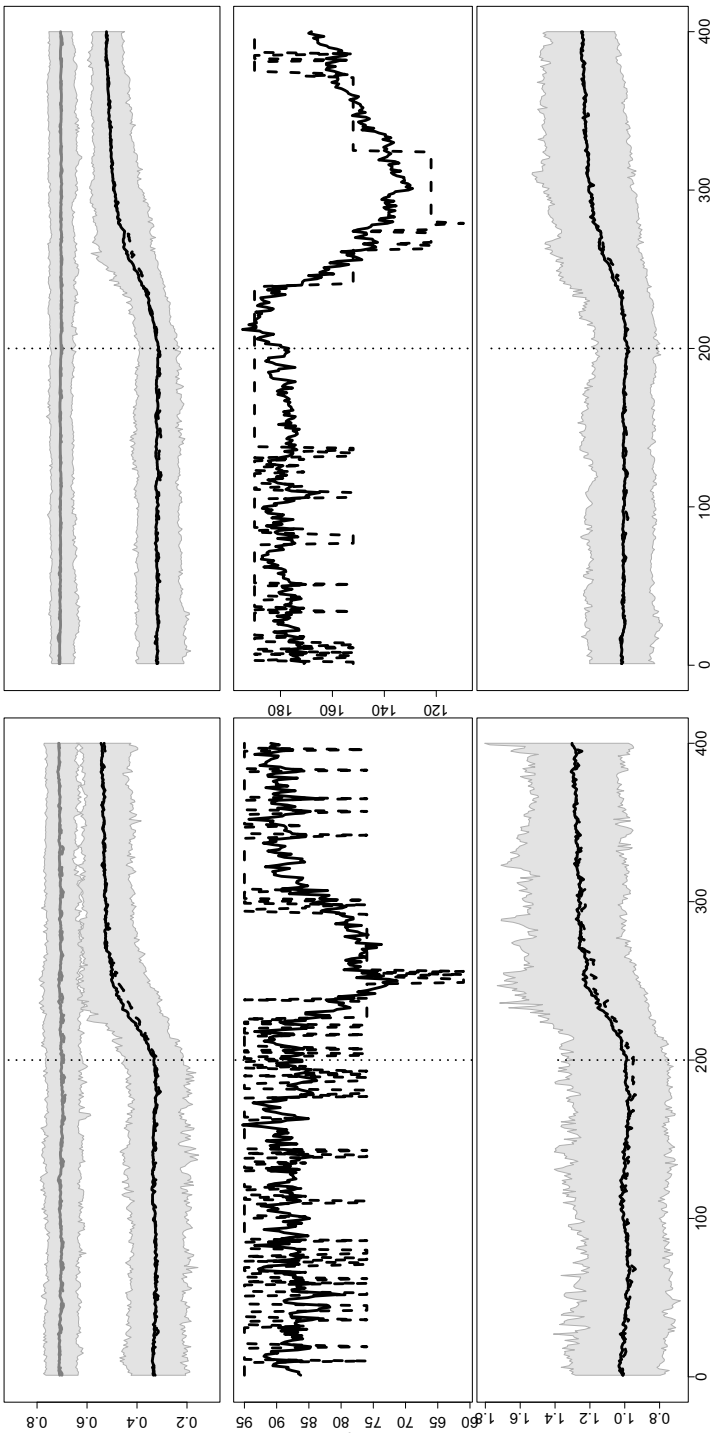


**Figure 4.3** The average estimated parameters, the length of the interval of homogeneity, and the ML criteria on the intervals of homogeneity for the model (4.1). The solid line depicts the average, the dashed line the median, and grey shaded area the 90% confidence interval for the respective variables over 100 replications. The shift occurs at observation 200. $m_0$ is set to 20 for the left panel and to 40 for the right panel, $\rho = 0.5$.

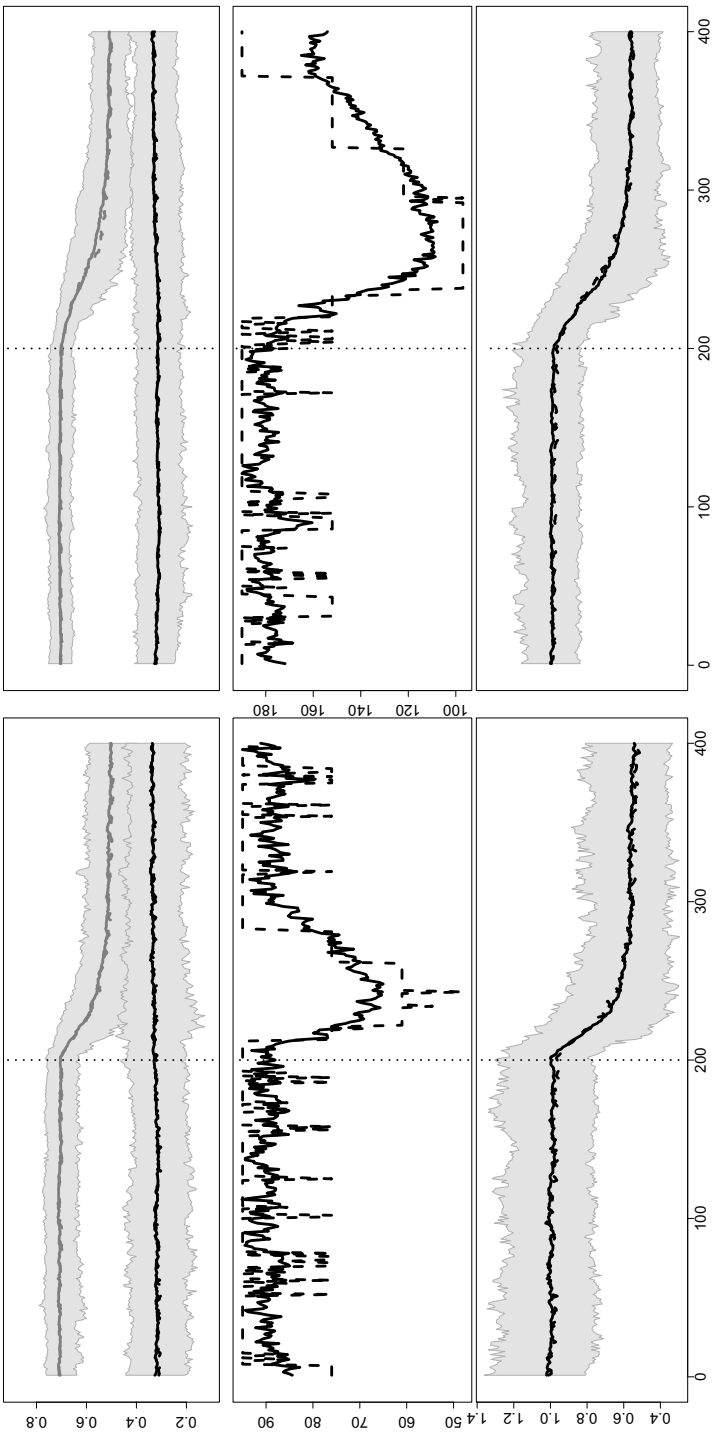


**Figure 4.4** The average estimated parameters, the length of the interval of homogeneity, and the ML criteria on the intervals of homogeneity for the model (4.2). The solid line depicts the average, the dashed line the median, and grey shaded area the 90% confidence interval for the respective variables over 100 replications. The shift occurs at observation 200. $m_0$ is set to 20 for the left panel and to 40 for the right panel, $\rho = 0.5$.

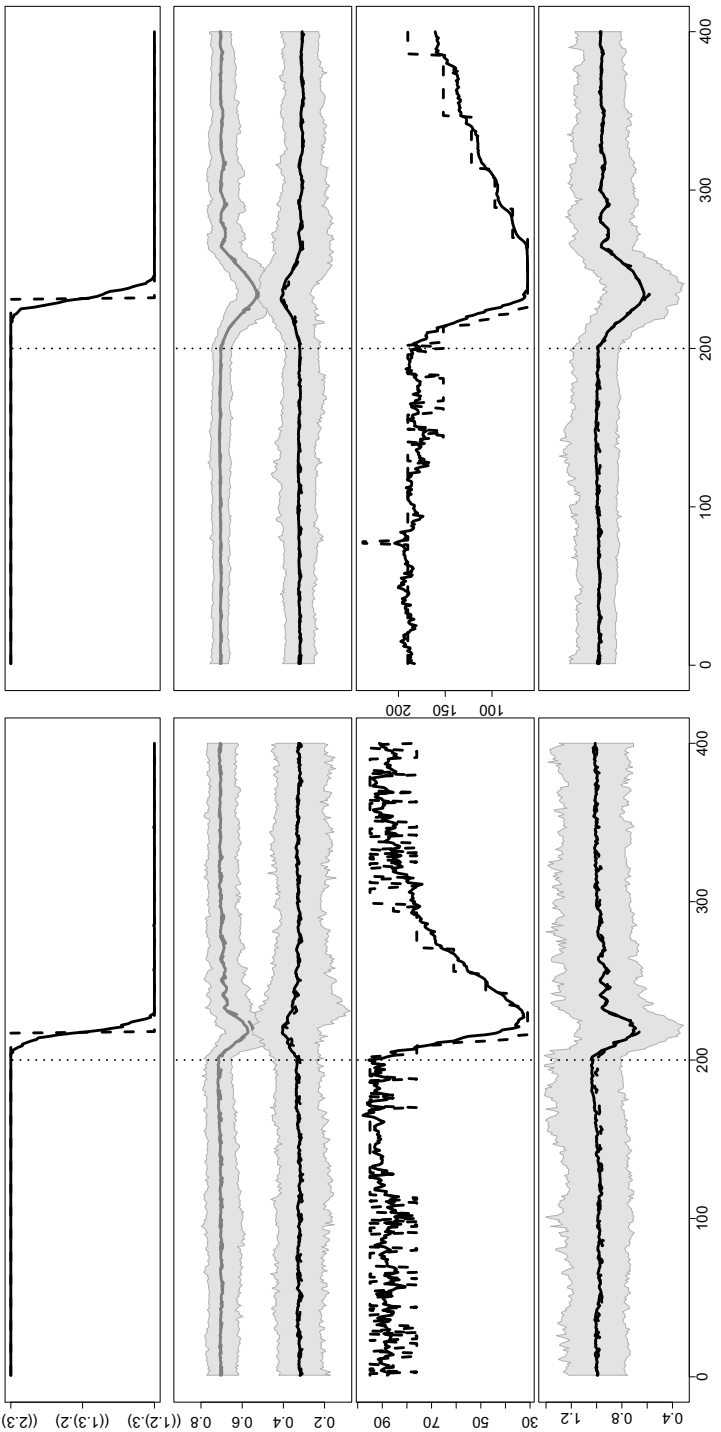


**Figure 4.5** The average estimated structure, parameters, the length of the interval of homogeneity, and the ML criteria on the intervals of homogeneity for the model (4.3). The solid line depicts the average, the dashed line the median and grey shaded area the 90% confidence interval for the respective variables over 100 replications. The shift occurs at observation 200. $m_0$ is set to 20 for the left panel and to 40 for the right panel, $\rho = 0.5$.

with $\theta(\tau_1, \tau_2)$ and the mis-specified copula with the same structure $s_0$ but with $\theta(0.1, 0.2)$. $\tau_1$ and $\tau_2$ take values between zero and one. Thus we observe in general an increase in the parameters from the true values $\theta(0.1; 0.2)$ to the mis-specified values $\theta(\tau_1; \tau_2)$. In the second term in (4.4) the situation is the opposite, and we observe a decrease in the copula parameters from the true values $\theta(\tau_1, \tau_2)$ to the mis-specified values $\theta(0.1, 0.2)$. The difference in (4.4) is plotted in Figure 4.2. The KL divergence is larger for increasing parameters and the difference becomes larger with increasing $\tau_1$ and $\tau_2$. This explains why the adaptive detection procedure based on the KL divergence reacts more quickly to an increase in parameters than to a decrease.

# 5 Empirical study

We now apply the local estimation procedure developed to multivariate data on stock indices and exchange rates. The data on indices contains the daily returns values for the Dow Jones (DJ), DAX, and NIKKEI, while the second data set consists of the daily values for the exchange rates JPN/USD, GBP/USD, and EUR/USD. Both data sets are taken from `DataStream`. The indices are obtained for the period [01.01.1985; 23.12.2010] resulting in 6778 observations, and the exchange rates cover the period [4.1.1999; 14.8.2009], resulting in 2771 observations.

To eliminate the intertemporal conditional heteroscedasticity we fit a univariate GARCH-type process to each marginal time series of log-returns. We assume that the corresponding parameters are constant over time, but select model specifications which reflect the peculiarities of the data such as asymmetries. Alternatively, the parameters of the marginal processes can be estimated by LCP as in Mercurio and Spokoiny (2004). Using the Bayes information criteria (BIC) as a goodness-of-fit measure, we selected for the indices the APARCH(1,1) model with the residuals following the skewed-$t$ distribution; i.e.,

$$X_{j,t} = \mu_j + \sigma_{j,t}\varepsilon_{j,t} \tag{5.1}$$

$$\text{with } \sigma_{j,t}^{\delta_j} = \omega_j + \upsilon_j\{|X_{j,t-1} - \mu_j| - \gamma(X_{j,t-1} - \mu_j)\}^{\delta_j} + \beta_j\sigma_{j,t-1}^{\delta_j}, \tag{5.2}$$

where $\varepsilon_{j,t} \sim t_{\text{skewed}}(\kappa; \nu)$, $j = 1, \ldots, 3$. The parameters $\kappa$ and $\nu$ stand for the skewness and shape (degrees of freedom) of the distribution. For the exchange rates, the best fit was a simple GARCH(1,1) process with GED residuals.

$$X_{j,t} = \mu_j + \sigma_{j,t}\varepsilon_{j,t} \text{ with } \sigma_{j,t}^2 = \omega_j + \upsilon_j\sigma_{j,t-1}^2 + \beta_j(X_{j,t-1} - \mu_j)^2 \tag{5.3}$$

and $\varepsilon_{j,t} \sim GED(\kappa_{GED}; \nu_{GED})$, $j = 1, \ldots, 3$. Similarly to the above, the parameters $\kappa_{GED}$ and $\nu_{GED}$ reflect the skewness and shape of the distribution.

The estimates of the parameters, $p$-values of the Box–Ljung test with 12 lags and the Kolmogorov–Smirnov test applied to the residuals, for the indices and exchange rates can be provided upon request. All the parameters are significant at the 5% significance level. The residuals exhibit the typical behaviour: they are not normally distributed, which motivates nonparametric estimation of the margins. From the results of the Box–Ljung test

we conclude that the autocorrelation of the residuals is significant only for the GBP/EUR rate. An additional autoregressive component, however, does not lead to a decrease in the BIC. Hereafter, we work only with the residuals of the fitted processes.

## 5.1 Rolling window estimation

The dependency variation is measured by Pearson's and Kendall's correlation coefficients: Figure 5.1 shows the behaviour of both coefficients calculated in a rolling window of width $r = 250$. Their dynamic behaviour is similar, but not identical. The shadowed area presents the 95% confidence intervals of the corresponding correlation measures. They allow us to conclude that the variation of the coefficients in time is statistically significant. This opens the door for a time varying copula based model.

To justify the use of a copula-based distribution to model the residuals, we further estimate alternative parametric models using a rolling window of the same width. We consider the binary HAC with Clayton or Gumbel generators; and the 3-dimensional Gaussian and 3-dimensional AC. The maximum-likelihood (ML) and the BIC are used as goodness-of-fit measures. Since the number of unknown parameters in the nonpara-

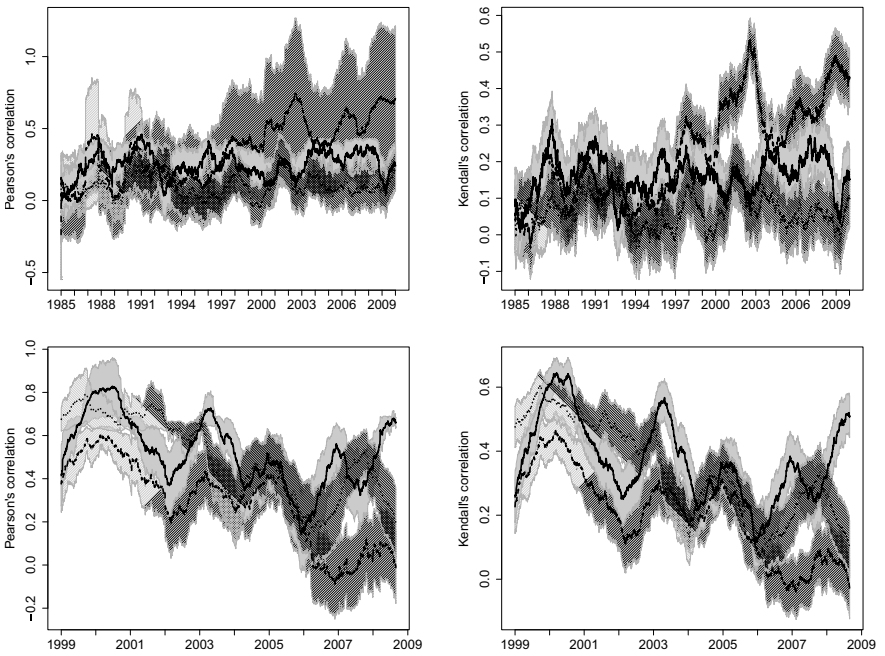


**Figure 5.1** Rolling window estimators of Pearson's (left) and Kendall's (right) correlation coefficients between the residuals of indices (top) and exchange rates (bottom). For the indices: DAX and NIKKEI (solid line), DAX and DJ (dashed line), DJ and NIKKEI (dotted line). For the exchange rates: JPY and EUR (solid line), JPY and GBP (dashed line), GBP and EUR (dotted line). The width of the rolling window is set to 250 observations. The shadowed area shows the 95% confidence interval around the corresponding estimates.

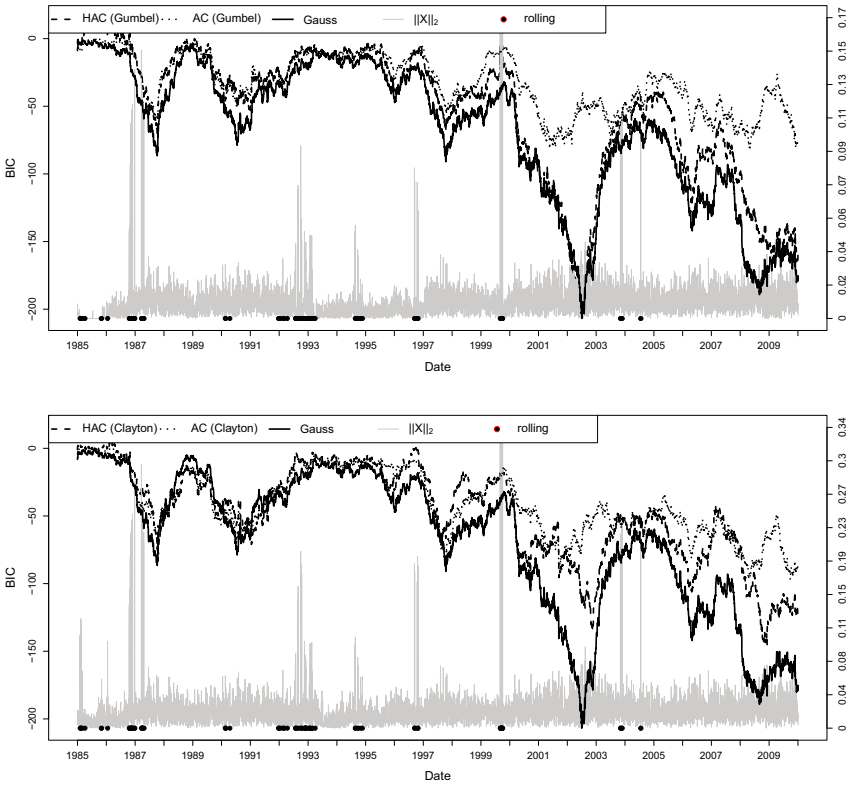


**Figure 5.2** BIC for indices over rolling window for HAC (dashed lines), AC (dotted lines) with Gumbel (top panel) and Clayton (bottom panel) generators. Solid line in both panels represents BIC for Gaussian copula. The width of the rolling window is 250 observations. The grey line shows the variation of $L_2$ norm of the difference in the parameter matrices of the copulae. The dots mark the changes in the structure of HAC using rolling window estimation.

metric case is unknown, it is incorrect to compare the models with nonparametrically and parametrically estimated margins using BIC. In such cases we consider only the parameters of the copula function. Figure 5.2 illustrates the dynamics of the BIC for five multivariate models for the indices. The HAC models is clearly dominated by the Gaussian model in some periods, but shows very competitive performance in other periods. Similar conclusions can be drawn for the exchange rates.

We check whether the variation in the dependency can be linked to some characteristics of the distribution. The dots in Figure 5.2 depict the time-points of changes in the HAC estimated using a rolling window procedure. There is no visible relationship between the dynamics of the model fit measured by the BIC and the changes in the structures. The thin grey line shows the dynamics of the $||\widehat{\boldsymbol{\Theta}}_t - \widehat{\boldsymbol{\Theta}}_{t-1}||_2$, where $\widehat{\boldsymbol{\Theta}}_t$ denotes the matrix of copula parameters estimated at $t$ and $||\cdot||_2$ denotes the $L_2$ matrix norm, which is defined

| Data | Generator | Structure | ML |
|---|---|---|---|
| Indices | Clayton | $((\text{DAX.DJ})_{0.45967(0.02125)}.\text{NIKKEI})_{0.15543(0.01226)}$ | 545.399 |
| | Gumbel | $((\text{DAX.DJ})_{1.27274(0.01255)}.\text{NIKKEI})_{1.10339(0.00765)}$ | 542.736 |
| Ex. rates | Clayton | $((\text{JPY.USD})_{0.80889(0.04261)}.\text{GBP})_{0.40134(0.02504)}$ | 617.268 |
| | Gumbel | $((\text{JPY.USD})_{1.52149(0.02504)}.\text{GBP})_{1.30340(0.01657)}$ | 736.341 |

**Table 5.1** Estimation results for HAC with Clayton and Gumbel generators for indices and exchange rates using full samples. The standard deviations of the parameters are given in parentheses.

by $||\mathbf{A}||_2 = \sqrt{\lambda_{\max}(\mathbf{A}^\top \mathbf{A})}$, where $\lambda_{\max}$ is the largest eigenvalue of the matrix $\mathbf{A}^\top \mathbf{A}$. HAC is closed under the marginalisation, meaning that taking any subset of variables following an HAC will also follow an HAC of smaller dimension, see Okhrin, Okhrin, and Schmid (2013b). This implies, that all the bivariate margins are AC with the parameters from the original HAC. For example, if $u_1, u_2, u_3$ follows $C_{\theta_1}\{C_{\theta_2}(u_1, u_2), u_3\}$, then $(u_1, u_2)$ follows $C_{\theta_2}(u_1, u_2)$, $(u_1, u_3)$ follows $C_{\theta_1}(u_1, u_3)$ and $(u_2, u_3)$ follows $C_{\theta_1}(u_1, u_2)$. Based on this we can define the matrix of parameters as $\mathbf{\Theta}_t = \{\theta_{i,j}\}_{i,j=1,\ldots,d;j\neq j}$, where $\theta_{i,j}$ corresponds to the parameter of the AC between $u_i$ and $u_j$. From the changes in the norms we see, that the parameters do contain important information about changes in the copula and their precise estimation achieved by the LCP procedure is of importance.

## 5.2 Local window estimation

The previous section provided evidence on two important issues. First, the univariate marginal distributions are not normal and the joint distribution can be modelled using an HAC based distribution. Second, the dependency is not constant, but varies with time. Since we model the dependency by HAC, this implies that either the structure of the HAC or the copula parameters are time-dependent. In this section we apply the local window procedure to compute a robust estimator of HAC.

The setup of our procedure is as follows. The set of $m_k$s defining the length of $I_k$ and $\mathcal{T}_k$ is determined by a geometric grid with $m_k = [m_0 c^k]$ for $k = 1, 2, \ldots, K$. The starting values are set to $m_0 = 40$, $\rho = 0.5$, and $c = 1.25$, where $[x]$ denotes the integer part of $x$. The critical values $\zeta$ are taken from the simulation study. This is a feasible procedure since the time series were filtered with the AR-GARCH processes to remove autocorrelation and heteroscedasticity. Thus the dependence structure of the residuals can be monitored using the critical values computed for the unautocorrelated uniformly distributed random variables. The structure, the parameters, and the corresponding standard deviations estimated from the whole data sample are given in Table 5.1.

Figures 5.3 and 5.4 illustrate the results of the application with Clayton generator. The upper plots show the changes in the structure. For the indices, the changes in structure are very frequent for the first half of the period, implying that the dependencies between different pairs of variables are similar. The structure ((1.2).3) is very robust in the second half of the period. This fact is supported by the rolling window estimation of the correlation coefficients in the previous section. Initially, the correlation lines intersect frequently. In this case, the procedure can hardly distinguish between different but similar structures.

In the second half of the period the correlation lines stay apart and a dominant structure evolves. In the analysis of exchange rates we observe two stable structures ((1.3).2) and ((2.3).1). On the other hand with the indices, the switches between the structures are not frequent.

The second two pictures show the parameters' estimators over the intervals of homogeneity transformed to Kendall's $\tau$. The grey line depicts the larger parameter, while the black line depicts the smaller parameter. If the order of the parameters changes, we observe a change in the structure. We see that the algorithm captures even relatively small changes in parameters. The figure for the indices shows that the parameters become more distant in the second half of the period. However, for the exchange rates, the parameters appear to be more widespread over the whole sample.

The third pictures indicate the dynamics of the ML criteria over the intervals of homogeneity. Recall that the local window procedure is based on the stability of the fit measured by maximum-likelihood. The overall fit of the HAC increases in time for the indices, but decreases for the exchange rates. Note that neither the changes in the structure nor changes in the parameters can explain the variation in the ML. However, the overlapping of both shifts closely follows the drift in the ML criteria. The bottom figures present the length of the intervals of homogeneity. For the indices, the intervals are close to the minimum of $m_0 = 40$ in the period of frequent changes in the structure, but steadily increases over stable periods. For the exchange rates, the intervals of homogeneity exhibit longer increases due to stabler structures and parameters. In general we observe clear variation in the lengths of the intervals of homogeneity, which justifies the application of the suggested methodology to HAC estimation.

Some changes in the structure or in the parameters can be caused by specific economic events. Regarding exchange rates, drastic changes in the parameters and some jumps in the structure at the beginning of 2000 reflect the fact that the euro reached its lowest level against the U.S. dollar in its history. Changes before the end of 2001 may be affected by the slumping US economy in August–September 2001, i.e., a negative growth rate in the third quarter 2001. The changes in the middle of 2003 are probably influenced by the rejection of the euro by Britain in June 2003. The changes in the structures and parameters in the model for indices can be partially explained by the following economic events. The jumps of the structure at the end of 1998 and the beginning of 1999 are caused by the financial crisis in Asia. The penultimate group of peaks in the changes of the structure were probably caused by the Iraq war. The last peaks in the structure plot, which also correspond to volatile behaviour in the parameters, corresponds to the changes of the government in Germany which influenced the DAX index, and a general election in Japan, influencing the NIKKEI.

## 5.3 Value-at-Risk with HAC

To assess the economic significance of the local estimation procedure for HAC, we consider the Value-at-Risk (VaR) of a portfolio of three assets with weights $\mathbf{w}$, assuming individual GARCH-type data generating processes for each asset. We fit an HAC with Gumbel and Clayton generators to each class of assets. The DCC model of Engle (2002) with Gaussian residuals serves as a benchmark. The profit and loss function of the

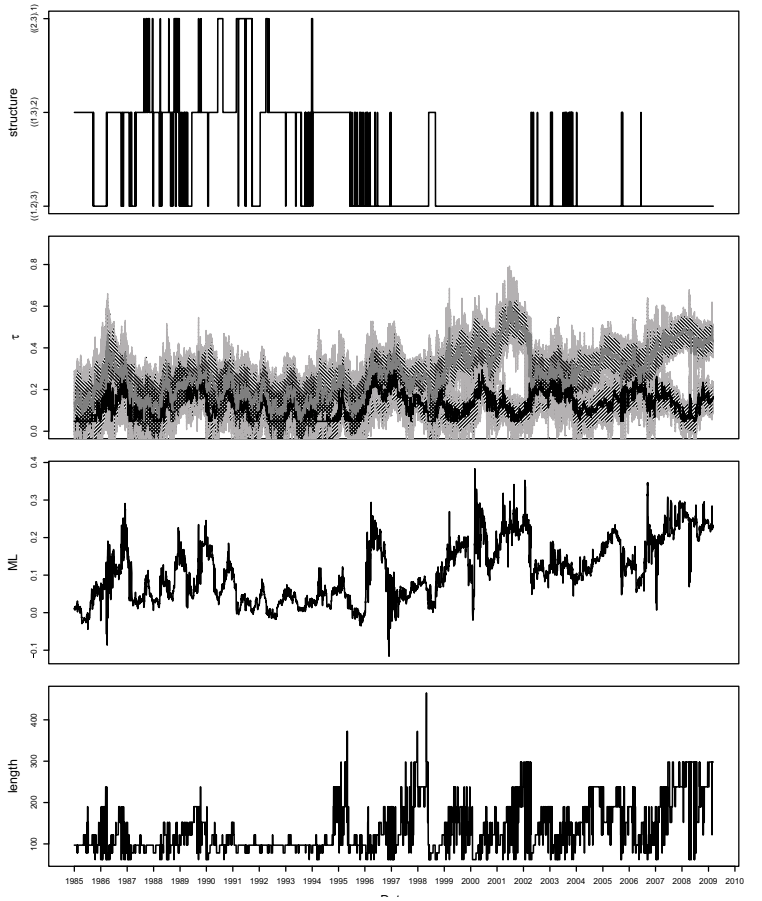


**Figure 5.3** Changes in the structure, changes in the parameters, and variation of the maximum-likelihood over the intervals of homogeneity for DJ, NIKKEI, and DAX modelled with binary Clayton HAC. $m_0$ is set to 40 and $\rho = 0.5$.

portfolio is defined as $L_{t+1} = \sum_{j=1}^{3} w_j P_{jt}(e^{R_{j,t+1}} - 1)$, where $P_{jt}$ and $R_{jt}$ are the price and the log-return of the asset $j$, respectively, at $t$. Let $F_L$ denote the distribution function of $L_{t+1}$. This leads to the VaR of the portfolio at level $\alpha$ being defined by $VaR(\alpha) = F_L^{-1}(\alpha)$. The distribution function $F_L$ is estimated by simulating the paths of the asset returns from the alternative multivariate processes. The $\widehat{\text{VaR}}(\alpha)$ is computed as the corresponding empirical quantile. Figure 5.5 shows the true path of the profit and loss function with the VaR estimator for equally weighted portfolio and the HAC with Clayton generator. The empirical PL function is shown with dots, while the VaR for different estimation techniques is shown by the solid line. The exceedances are marked with pluses on the horizontal axis .

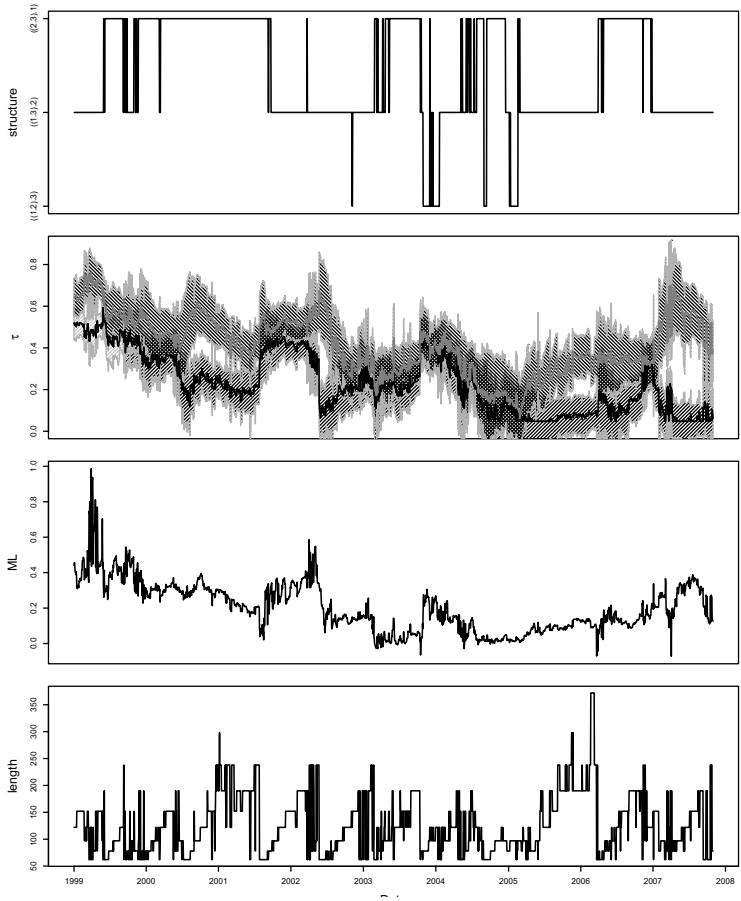


**Figure 5.4** Changes in the structure, changes in the parameters, and variation of the maximum-likelihood over the intervals of homogeneity for JPY, GBP, EUR modelled with binary Clayton HAC. $m_0$ is set to 40 and $\rho = 0.5$.

Using backtesting, we assess the economic significance of the chosen model for the assets. We estimate the realised $\alpha$ as the relative fraction of the exceedances in the time series, i.e.,

$$\hat{\alpha}_{\mathbf{w}} = \frac{1}{T}\sum_{t=1}^{T} \mathbf{I}\{L_t < \widehat{\mathrm{VaR}}_t(\alpha)\}.$$

The precision of the underlying model is measured by the relative distance between the estimated $\hat{\alpha}_{\mathbf{w}}$ and the true $\alpha$, calculated by

$$e_{\mathbf{w}} = (\hat{\alpha}_{\mathbf{w}} - \alpha)/\alpha.$$

| | | Clayton | | | Gumbel | | | DCC | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | $\alpha$ | 0.0100 | 0.0500 | 0.1000 | 0.0100 | 0.0500 | 0.1000 | 0.0100 | 0.0500 | 0.1000 |
| Indices | $\hat{\alpha}_{w^*}$ | 0.0054 | 0.0390 | 0.0935 | 0.0033 | 0.0249 | 0.0683 | 0.0155 | 0.0460 | 0.0830 |
| | $A_W$ | −0.3902 | −0.1781 | −0.0497 | −0.6187 | −0.4496 | −0.2686 | 0.5979 | −0.0687 | −0.1739 |
| | $D_W$ | 0.0930 | 0.0508 | 0.0286 | 0.0953 | 0.0932 | 0.0638 | 0.0959 | 0.0829 | 0.0609 |
| Ex. Rates | $\hat{\alpha}_{w^*}$ | 0.0100 | 0.0487 | 0.0951 | 0.0091 | 0.0474 | 0.0977 | 0.0156 | 0.0413 | 0.0817 |
| | $A_W$ | −0.0217 | −0.0328 | −0.0557 | −0.0895 | −0.0526 | −0.0341 | 0.5482 | −0.1652 | −0.1852 |
| | $D_W$ | 0.0649 | 0.0186 | 0.0125 | 0.0632 | 0.0406 | 0.0272 | 0.0335 | 0.0091 | 0.0042 |

**Table 5.2** Exceedance ratios for portfolios of indices (top) and exchange rates (bottom) with $w^*$, $w_i, i = 1, \ldots, 5$, the average exceedance $A_W$ over all portfolios, and its standard deviation $D_W$.

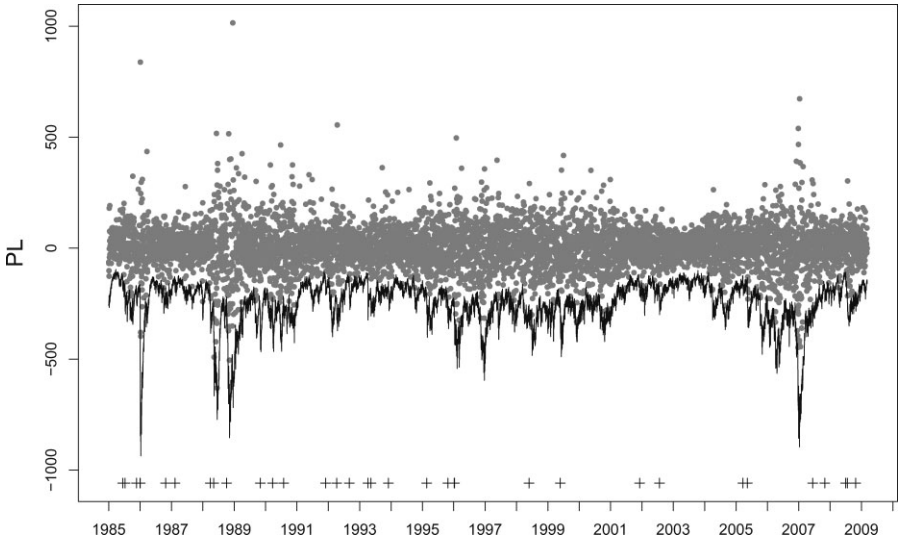


**Figure 5.5** The profit and loss function (dots) for indices with the 99%-VaR bound (solid line) and the time points of the exceedances (pluses) for the HAC model with Clayton generator.

As in Giacomini, Härdle, and Spokoiny (2009), we compute $\hat{\alpha}_{\mathbf{w}}$ and $e_{\mathbf{w}}$ for a set $W = \{\mathbf{w}^*, \mathbf{w}_n; n = 1, \ldots, 99\}$ of portfolios: for each $\mathbf{w}_n = (w_{n,1}, w_{n,2}, w_{n,3})'$ is the realisation of a random vector uniformly distributed on $S = \{(x_1, x_2, x_3) \in \mathbb{R}^3 : \sum_{i=1}^3 x_i = 1, x_i \geq 0.1\}$ and $\mathbf{w}^* = (1/3, 1/3, 1/3)'$ is the equally weighted portfolio. The performance of each model is measured by the average relative exceedance over the portfolios and its corresponding standard deviation

$$A_W = \frac{1}{|W|} \sum_{\mathbf{w} \in W} e_{\mathbf{w}}, \qquad D_W = \left\{ \frac{1}{|W|} \sum_{\mathbf{w} \in W} (e_{\mathbf{w}} - A_W)^2 \right\}^{1/2}.$$

The results of this backtesting are summarised in Table 5.2 for indices and exchange rates. For the indices, we conclude that the Clayton-based HAC outperforms the HAC

with Gumbel generator and the DCC model for small and large levels of $\alpha$, while DCC shows smaller $A_W$ for $\alpha = 0.05$. Moreover, the copula-based models tend to underestimate the true level, while DCC overestimates it for $\alpha = 0.01$. For the exchange rates, the HAC with both generators is clearly dominant compared with the DCC model, and the general performance of the copula-based models is significantly better than for the indices.

# 6 Conclusions

We proposed a method of estimating time-varying dependencies. The joint distribution of multivariate observations is modelled by a Hierarchical Archimedean copula characterised by a dependency structure and generator function. We assume that over specific time intervals, the time-varying copulae can be approximated by constant copulae. The optimal intervals for these homogeneous structures and homogeneous parameters was determined using a Local Change Point detection procedure. This proposal was evaluated in an extensive simulation study and compared to the classical rolling window estimation. The optimal estimation period rapidly drops after a shift in the structure or in the parameters, allowing for less biased estimators. The real data application was performed using a three dimensional time series of index returns and exchange rates. The results disclosed periods of stable structure in both data sets and strongly varying periods of homogeneity, leading to the conclusion that the rolling window estimation cannot be suitable for the data. The economic significance of the suggested estimation procedure was evaluated with VaR for portfolios. The locally estimated HAC clearly dominated the popular DCC model used as a benchmark. Summarising, the local estimation procedure improves the properties of the estimators and is economically attractive for modelling time-varying dependencies of financial assets.

# References

Baba, Y., Engle, R., Kraft, D. and Kroner, K. (1990). Multivariate simultaneous generalized ARCH, *Technical Report (unpublished manuscript)*, University of California, San Diego.

Breymann, W., Dias, A. and Embrechts, P. (2003). Dependence structures for multivariate high-frequency data in finance, *Quantitative Finance* **1**: 1–14.

Chen, X. and Fan, Y. (2006). Estimation and model selection of semiparametric copula-based multivariate dynamic models under copula misspesification, *Journal of Econometrics* **135**: 125–154.

Chen, Y., Härdle, W. and Jeong, S.-O. (2008). Nonparametric risk management with generalized hyperbolic distributions, *Journal of the American Statistical Association* **103**(483): 910–923.

Embrechts, P., Lindskog, F. and McNeil, A. J. (2003). Modeling dependence with copulas and applications to risk management, *in* S. T. Rachev (ed.), *Handbook of Heavy Tailed Distributions in Finance*, Amsterdam, Elsevier.

Engle, R. (2002). Dynamic conditional correlation – a simple class of multivariate GARCH models, *Journal of Business and Economic Statistics* **20**(3): 339–350.

Giacomini, E., Härdle, W. K. and Spokoiny, V. (2009). Inhomogeneous dependence modeling with time-varying copulae, *Journal of Business and Economic Statistics* **27**(2): 224–234.

Hering, C., Hofert, M., Mai, J.-F. and Scherer, M. (2010). Constructing nested Archimedean copulas with Lévy subordinators, *Journal of Multivariate Analysis* **101**(6): 1428–1433.

Hofert, M. (2012). A stochatic representation and sampling algorithm for nested Archimedean copulas, *Journal of Statistical Computation and Simulation* **82**(9): 1239–1255.

Joe, H. (1996). Families of $m$-variate distributions with given margins and $m(m-1)/2$ bivariate dependence parameters, *in* L. Rüschendorf, B. Schweizer, and M. Taylor (eds.), *Distribution with Fixed Marginals and Related Topics*, IMS Lecture Notes – Monograph Series, Institute of Mathematical Statistics.

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

Jondeau, E. and Rockinger, M. (2006). The copula-GARCH model of conditional dependencies: An international stock market application, *Journal of International Money and Finance* **25**(5): 827–853.

Lamoureux, C. G. and Lastrapes, W. D. (1990). Persistence-in-variance, structural change and the GARCH model, *Journal of Business and Economic Statistics* **8**: 225–234.

McNeil, A. J. (2008). Sampling nested Archimedean copulas, *Journal Statistical Computation and Simulation* **78**(6): 567–581.

McNeil, A. J. and Nešlehová, J. (2009). Multivariate Archimedean copulas, $d$-monotone functions and $l_1$ norm symmetric distributions, *Annals of Statistics* **37**(5b): 3059–3097.

Mercurio, D. and Spokoiny, V. (2004). Statistical inference for time-inhomogeneous volatility models, *Annals of Statistics* **32**(2): 577–602.

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

Okhrin, O., Okhrin, Y. and Schmid, W. (2013a). On the structure and estimation of hierarchical Archimedean copulas, *Journal of Econometrics* **173**(2): 189–204.

Okhrin, O., Okhrin, Y. and Schmid, W. (2013b). Properties of Hierarchical Archimedean Copulas, *Journal of Risk and Modeling* **30**(1): 21–53.

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

Rodriguez, J. C. (2007). Measuring financial contagion: A copula approach, *Journal of Empirical Finance* **14**: 401–423.

Savu, C. and Trede, M. (2010). Hierarchies of archimedean copulas, *Quantitative Finance* **10**: 295–304.

Schmid, F. and R. Schmidt (2006). Multivariate extensions of Spearman's rho and related statistics, *Statistics and Probability Letters* **77**(4): 407–416.

Silvennoinen, A. and Teräsvirta, T. (2009). Multivariate GARCH models, *in* T. G. Andersen, R. A. Davis, J.-P. Kreiß, and T. Mikosch (eds.), *Handbook of Financial Time Series*, Springer-Verlag, Berlin, pp. 201–233.

Sklar, A. (1959). Fonctions de repartition á *n* dimension et leurs marges, *Publ. Inst. Stat. Univ. Paris* **8**: 299–231.

Spokoiny, V. (2009). Multiscale local change point detection with applications to value-at-risk, *The Annals of Statistics* **37**(3): 1405–1436.

Spokoiny, V. (2010). *Local Parametric Methods in Nonparametric Estimation*, Springer-Verlag, Berlin. forthcoming.

Čížek, P., Härdle, W. and Spokoiny, V. (2009). Adaptive pointwise estimation in time-inhomogeneous conditional heteroscedasticity models, *Econometrics Journal* **12**(2): 248–271.

Whelan, N. (2004). Sampling from Archimedean copulas, *Quantitative Finance* **4**: 339–352.

White, H. (1982). Maximum likelihood estimation of misspecified models, *Econometrica* **50**: 1–25.

Wolfgang Karl Härdle
C.A.S.E.
Center for Applied Statistics and Economics
Humboldt-Universität zu Berlin
Spandauer Straße 1
10178 Berlin
Germany
haerdle@wiwi.hu-berlin.de.

Ostap Okhrin
C.A.S.E.
Center for Applied Statistics and Economics
Humboldt-Universität zu Berlin
Spandauer Straße 1
10178 Berlin
Germany
ostap.okhrin@wiwi.hu-berlin.de

Yarema Okhrin
Department of Statistics
University of Augsburg
86135 Augsburg
Germany
yarema.okhrin@wiwi.uni-augsburg.de