

# Efficient Bayesian inference for stochastic time-varying copula models

Carlos Almeida<sup>a,\*</sup>, Claudia Czado<sup>a</sup>

<sup>a</sup>*Department of Mathematical Statistics, Technische Universität München, Germany*

---

## Abstract

There is strong empirical evidence that dependence in multivariate financial time series varies over time. To incorporate this effect we suggest a time varying copula class, which allows for stochastic autoregressive (SCAR) copula time dependence. For this we introduce latent variables which are analytically related to Kendall's  $\tau$ , specifically we introduce latent variables that are the Fisher transformation of Kendall's  $\tau$  allowing for easy comparison of different copula families such as the Gaussian, Clayton and Gumbel copula. The inclusion of latent variables renders maximum likelihood estimation computationally infeasible, therefore a Bayesian approach is followed. Such an approach also enables credibility intervals to be easily computed in addition to point estimates. We design two sampling approaches in a Markov Chain Monte Carlo (MCMC) framework. The first is a naïve approach based on Metropolis-Hastings in Gibbs while the second is a more efficient coarse grid sampler using ideas of Liu and Sabatti (2000). The performance of these samplers are investigated in a large simulation study and are applied to two data sets involving financial stock indices. It is shown that time varying dependence is present for these data sets and can be quantified by estimating time varying Kendall's  $\tau$  with point-wise credible intervals over the series.

*Keywords:* time varying dependence, non Gaussian copulas, Kendall's  $\tau$ , Bayesian inference, Markov Chain Monte Carlo

---

\*Corresponding author: Carlos Almeida, Department of Mathematical Statistics, Technische Universität München, Boltzmannstr. 3, D-85747 Garching, Germany, phone: +49 89 28917439

*Email addresses:* calmeida@ma.tum.de (Carlos Almeida), czado@ma.tum.de (Claudia Czado)

## 1. Introduction

Since the introduction of copulas by Sklar (1959) as tool for constructing multivariate distributions they have become and its increasingly popular in a variety of fields (see for example the books by Joe (1997) and Nelsen (2006)). One such field is finance, where they are used to determine Value-at-Risk of portfolios, to construct optimal portfolios and for pricing financial products with several underlying assets. The book by Cherubini, Luciano, and Vecchiato (2004) is dedicated to various applications of copulas to finance.

Copulas also play an important role in multivariate GARCH models (see Bauwens, Laurent, and Rombouts (2006) for a survey of such models) suitable for modeling financial time series. Initially such models allowed only for time constant dependencies, however empirical work shows that this assumption is not suitable for many data sets; see for example Erb, Harvey, and Viskanta (1994), Longin and Solnik (1995) and Engle (2002). This insight started a strong interest in copula based models, which allow for time varying dependence parameters. Dias and Embrechts (2004) and Manner and Candelon (2007) use a change point approach to identify a change in copula dependence, while Giacomini, Härdle, and Spokoiny (2009) use to the local change point (LCP) method of Mercurio and Spokoiny (2004).

A recent survey of time varying copula models is given by Manner and Reznikova (2009). Early time-varying dependence models are the DCC models proposed by Tse and Tsui (2002) and Engle (2002), which model conditional correlations. They are observation driven and require special efforts to achieve a positive definite correlation matrix. As noted by Bauwens, Laurent, and Rombouts (2006) a drawback of these DCC models is that the parameters needed to model time dependencies are scalar and thus imply that the conditional correlations between pairs of variables obey the same nonstochastic dynamics. Additionally, since these models are correlation based they incorporate only dependencies allowed by elliptical distributions.

More general are copula-GARCH models which were suggested by Patton (2006) and Jondeau and Rockinger (2006). These models assume GARCH margins, while the dependence is modelled by a copula. Estimation is usually earned out in a two step approach, since a joint estimation of all parameters is too costly. First the marginal parameters are estimated and then these estimates are used to transform the standardized innovations via the probability integral transform to copula data. Finally this derived copula data is used to estimate the copula parameters. Joe (2005) and Patton (2006)

have shown that this leads to consistent, but not efficient estimators, when maximum likelihood (ML) estimators are used at the two estimation steps. Consistency is also achieved if one uses normalized ranks of the standardized marginal innovations to transform to copula data. Liu and Luger (2009) have proposed an algorithm to improve the efficiency in copula-GARCH models when ML estimation is used. Subsequent time-varying copula-GARCH models have been proposed by Jondeau and Rockinger (2006) and Patton (2006).

As an alternative to GARCH models, stochastic volatility (SV) models (see for example Taylor (1986)) can be used as marginal models together with a copula to construct multivariate financial time series models. Multivariate generalizations of the SV are proposed by Harvey, Ruiz, and Shephard (1994) while Yu and Meyer (2006) consider a bivariate SV model with time dependent stochastic correlations assuming multivariate normal or  $t$  errors. Assuming marginal SV models together with an arbitrary bivariate copula family has been considered by Hafner and Manner (2008). They allow for time-varying stochastic copula parameters by choosing a transformation from the parameter space of the copula parameter to the real numbers and assuming a Gaussian AR(1) model for the transformed copula parameter. Estimation is facilitated by using efficient importance sampling (Liesenfeld and Richard (2003) and Richard and Zhang (2007)), however standard error estimates are difficult to compute.

In a Bayesian approach using Markov Chain Monte Carlo (MCMC) algorithms precision, estimates are easy to obtain, since samples of the posterior are available to construct credible intervals. The advantage of the Bayesian approach has been recognized in this context by Yu and Meyer (2006) and Ausin and Lopes (2009). However it is vital to construct and implement MCMC algorithms that are fast and mix well. Yu and Meyer (2006) use WINBUGS, which utilizes individual Gibbs sampling for each component of the parameter vector to be estimated. They contend that this might not be an efficient way of sampling from the posterior. In contrast Ausin and Lopes (2009) considers a copula-GARCH model with the nonstochastic dependence dynamics chosen as in Tse and Tsui (2002). This restricts the class of copulas considered to the class of elliptical copulas. Ausin and Lopes (2009) estimate marginal and copula parameters jointly. They use joint multivariate random walk Metropolis-Hastings (MH) updates for the dynamic copula parameters and the GARCH parameters of each margins, respectively. They report that individual random walk MH-updates result in slow mixing of the MCMC samples. We like to note that the dynamics of the copula parameters con-

sidered in Ausin and Lopes (2009) are nonstochastic and observation driven Bayesian MCMC algorithms are much simpler.

The purpose of this paper is to develop efficient MCMC estimation algorithms for stochastic time-varying copula models. As stochastic dynamics we consider a Gaussian AR(1) model for the inverse Fisher transformation of the Kendall's  $\tau$  parameter corresponding to the chosen copula. This is in line with the model considered in Hafner and Manner (2008). Our approach is valid for any copula specified by a single parameter, and we specifically consider Gaussian, (double) Clayton and the (double) Gumbel copula. We follow a latent variable approach based on the *data augmentation principle* of Tanner and Wong (1987). A first naïve Gibbs sampling method for updating the latent variables individually is developed. As expected this is not very efficient and we improve the sampler by developing by a coarse grid sampler as introduced by Liu and Sabatti (2000). The model is applied to two data sets and interesting empirical results are uncovered.

The paper is organized as follows. In Section 2 we introduce the stochastic dynamic copula model and choose appropriate priors. In Section 3 we derive expressions for the full conditional densities which are needed for a Metropolis-Hastings within Gibbs sampler. An appropriate coarse grid method for updating the latent variables is developed. A large simulation study to investigate the behaviour of the MCMC samplers is conducted in Section 4. Two applications to financial stock indices are presented in Section 5. Finally, concluding a summary and an outlook are given in Section 6.

## 2. Model

For the general setup we use the following notations. Let  $(y_1, y_2) \in \mathbb{R}^2$  be a bivariate random vector with marginal cumulative distribution functions (cdf)  $F_i, i = 1, 2$  and joint cdf  $H$ . Using Sklar's theorem (Sklar (1959)), we can express  $H$  as  $H(y_1, y_2) = C(F_1(y_1), F_2(y_2))$ , where  $C$  is a copula cdf. For absolutely continuous distributions this can be rewritten for densities as  $h(y_1, y_2) = c(F_1(y_1), F_2(y_2))f_1(F_1(y_1))f_2(F_2(y_2))$ , where  $c$  is the corresponding copula density.

In financial applications knowledge of the past is collected in the filtration  $\mathcal{F}_{t-1}$  and we observe time series data  $(y_{1,t}, y_{2,t})$  for  $t = 1, \dots, T$ . Patton (2006) models the conditional joint distribution of  $(y_{1,t}, y_{2,t})$  given  $\mathcal{F}_{t-1}$  as

$$H_t(y_{1,t}, y_{2,t} \mid \mathcal{F}_{t-1}) = C(F_{1,t}(y_{1,t} \mid Z_{t-1}, \boldsymbol{\beta}_1), F_{2,t}(y_{2,t} \mid Z_{t-1}, \boldsymbol{\beta}_2) \mid Z_{t-1}, \boldsymbol{\alpha}).$$

Here  $Z_{t-1}$  is a set of variables in  $\mathcal{F}_{t-1}$ , while  $\beta_i, i = 1, 2$  are parameters for the marginal models and  $\alpha$  for the dependence model, respectively. It follows immediately that the log-likelihood ( $\mathcal{L}$ ) can be decomposed into the following form:

$$\mathcal{L}_{12}(\beta_1, \beta_2, \alpha) = \mathcal{L}_1(\beta_1) + \mathcal{L}_2(\beta_2) + \mathcal{L}_C(\beta_1, \beta_2, \alpha)$$

which separates the contributions to the joint likelihood ( $\mathcal{L}_{12}$ ) into marginal contributions ( $\mathcal{L}_i, i = 1, 2$ ) and the copula contribution ( $\mathcal{L}_C$ ). This particular form motivates a two stage estimator.

In the first stage the marginal parameters  $\beta_i, i = 1, 2$  are estimated by maximizing over  $\mathcal{L}_i(\beta_i)$  separately, giving  $\hat{\beta}_i$ . For the second stage the dependence parameter  $\alpha$  is found by maximizing  $\mathcal{L}_C(\beta_1, \hat{\beta}_2, \alpha)$ . This approach is called *inference function for margins* or IFM method. It was introduced by Joe (1996) and who proved later in Joe (1997) the consistency of this method.

Genest, Ghoudi, and Rivest (1995) proposes another two step estimator, where the margins are transformed by ranks to copula data in the first stage. This corresponds to an empirical probability transformation. In the second stage only the dependence parameter  $\alpha$  is estimated. This semi-parametric method is also consistent and more robust so long as misspecification of margins is not severe (see Kim, Silvapulle, and Silvapulle (2007)). Consequently, if the marginal models are carefully chosen to provide good marginal fitting, we can concentrate on estimating the dependence model using the marginally fitted parameters, or we could apply the empirical transformation if we are unsure of the marginal models. Thus we concentrate on building a copula model on  $[0, 1]^2$  which allows for time varying dependence.

Let  $\mathbf{u}_t = (u_{t1}, u_{t2})^\top \in [0, 1]^2: t = 1, \dots, T$  be a sample such a that:

$$\mathbf{u}_t \mid (\mathbf{u}_1, \dots, \mathbf{u}_{t-1}), (\boldsymbol{\theta}_1, \dots, \boldsymbol{\theta}_t) \sim C_{\boldsymbol{\theta}_t}(\cdot), \quad (1)$$

where  $C_{\boldsymbol{\theta}_t}(\cdot)$  stands for a parametric copula distribution with time varying parameter  $\boldsymbol{\theta}_t \in \Theta \subset \mathbb{R}^M$ . The parameters  $\boldsymbol{\theta}_t$  will be modeled as latent variables with the following dependence structure. The latent variable  $\boldsymbol{\theta}_{t-1}$  influences directly only  $u_{t-1}$  and  $\boldsymbol{\theta}_t$  as shown in Figure 1.

— Figure 1 about here —

This time dependence structure implies the following conditional independence conditions:

1. The variables  $\mathbf{u}_1, \dots, \mathbf{u}_T$  are conditionally independent given the latent variables  $\boldsymbol{\theta}_1, \dots, \boldsymbol{\theta}_{t-1}$ . In the notation from Florens, Mouchart, and Rolin (1990), this can be written as  $\perp\!\!\!\perp_{1 \leq t \leq T} \mathbf{u}_t \mid \boldsymbol{\theta}_1, \dots, \boldsymbol{\theta}_T$ , or equivalently in densities:

$$p(\mathbf{u}_1, \dots, \mathbf{u}_T \mid \boldsymbol{\theta}_1, \dots, \boldsymbol{\theta}_T) = \prod_{t=1}^T p(\mathbf{u}_t \mid \boldsymbol{\theta}_1, \dots, \boldsymbol{\theta}_T) \quad (2)$$

2. The present observation  $\mathbf{u}_t$  does not depend on the past latent variables  $\boldsymbol{\theta}_1, \dots, \boldsymbol{\theta}_{t-1}$  given the present value of the latent variable  $\boldsymbol{\theta}_t$ , *i.e.*  $\mathbf{u}_t \perp\!\!\!\perp \boldsymbol{\theta}_1, \dots, \boldsymbol{\theta}_{t-1} \mid \boldsymbol{\theta}_t$ , in densities:

$$p(\mathbf{u}_t \mid \boldsymbol{\theta}_1, \dots, \boldsymbol{\theta}_t) = p(\mathbf{u}_t \mid \boldsymbol{\theta}_t). \quad (3)$$

Although the parametric bivariate copulas involved can be general, for illustration purposes, we restrict to families governed by a single parameter, *i.e.*  $M = 1$ , or equivalently the parameter space  $\Theta \subset \mathbb{R}$ .

The fact that different parametric copula families have different parameter spaces, motivates us to transform these parameter spaces in a way that different families can share the same domain. From the literature on copulas, *e.g.* Joe (1997), Nelsen (2006), or Genest and Favre (2007), a generally accepted measure of dependence is the Kendall's  $\tau$  coefficient. We propose to model the time dependence of pairs of variables by allowing a time varying Kendall's  $\tau_t$  for the copula  $C_{\boldsymbol{\theta}_t}(\cdot)$ . For this note that most bivariate copula families have a one-to-one transformation from the parameter space  $\Theta$  into the range of values for the corresponding Kendall's  $\tau$ . We denote by  $\tau_t := \tau(\boldsymbol{\theta}_t)$  this relationship between the parameter specifying the copula and the corresponding Kendall's  $\tau_t$ . Then we consider a one-to-one transformation from  $\mathbb{R}$  into  $(-1, 1)$ , the range of the Kendall's  $\tau_t$  values. We denote this bijection as  $\gamma_t \mapsto \tau_t := h(\gamma_t)$ . For the simulation study and applications we utilise the inverse of the Fisher transformation as a bijection defined as:

$$\tau_t = \frac{\exp(2\gamma_t) - 1}{\exp(2\gamma_t) + 1} =: h(\gamma_t).$$

Thus, the relationship between the latent variables  $\gamma_t$  and the time varying parameters  $\boldsymbol{\theta}_t$  governing the bivariate copulas is given by the transformation:

$$\boldsymbol{\theta}_t \in \Theta \mapsto \gamma_t = h^{-1}(\tau(\boldsymbol{\theta}_t)) = \frac{1}{2} \log \left\{ \frac{1 + \tau(\boldsymbol{\theta}_t)}{1 - \tau(\boldsymbol{\theta}_t)} \right\} \in \mathbb{R}. \quad (4)$$

By assuming a time series model for  $\gamma_t$ , we also assume stochastic time varying copula parameters  $\theta_t$ , specifically, we assume  $\gamma_t$  follows a stationary AR(1) model *i.e.*:

$$\begin{aligned}\gamma_t &= \mu + \phi(\gamma_{t-1} - \mu) + \sigma\varepsilon_t \\ \varepsilon_t &\sim N(0, 1), \text{ i.i.d. } \forall t.\end{aligned}\tag{5}$$

Here  $N(\mu, \sigma^2)$  denotes a normal distribution with mean  $\mu$  and variance  $\sigma^2$ . We have chosen to model the time varying dependence using time varying  $\gamma_t$  which are analytically related to Kendall's  $\tau_t$  and not to the copula specific parameters  $\theta_t$ , as thus allows for easy comparison of different copula families. In addition Kendall's  $\tau$  can easily be empirically estimated.

For this model the copula dependence parameter vector is therefore given by  $\boldsymbol{\alpha} := (\mu, \phi, \sigma) \in \mathbb{R} \times [-1, 1] \times \mathbb{R}_+$ . We call the model specified in (1) and (5) a stochastic copula autoregressive (SCAR) model as proposed by Hafner and Manner (2008). We use the abbreviation  $SCAR(\mu, \phi, \sigma)$  in the subsequent text.

*Likelihood.* In the following we use the notation:  $\mathbf{x}_{1:N} := (x_1, \dots, x_N)^\top$ . To specify the likelihood, we apply the following conditional independence conditions:

$$\begin{aligned}(a) \quad & \perp\!\!\!\perp_{1 \leq t \leq T} \mathbf{u}_t \mid \boldsymbol{\gamma}_{1:T} \\ (b) \quad & \mathbf{u}_{1:T} \perp\!\!\!\perp \boldsymbol{\alpha} \mid \boldsymbol{\gamma}_{1:T} \\ (c) \quad & \mathbf{u}_t \perp\!\!\!\perp \boldsymbol{\gamma}_{1:T}, \boldsymbol{\alpha} \mid \gamma_t \\ (d) \quad & \gamma_t \perp\!\!\!\perp \boldsymbol{\gamma}_{1:(t-1)} \mid \gamma_{t-1}, \boldsymbol{\alpha}.\end{aligned}\tag{6}$$

The conditions (a), (b) and (c) in equation (6) formalise the idea that the dependence across time is given only through the latent variables as shown in Figure 1. The condition (d) specifies the time dependence by using the Markov property of the underlying AR(1) model.

Therefore the joint likelihood of the observations  $\mathbf{u}_{1:T}$  and the latent variables  $\boldsymbol{\gamma}_{1:T}$  is given by:

$$p(\mathbf{u}_{1:T}, \boldsymbol{\gamma}_{1:T} \mid \boldsymbol{\alpha}) = \prod_{1 \leq t \leq T} p(\mathbf{u}_t \mid \gamma_t) \cdot p(\gamma_1 \mid \boldsymbol{\alpha}) \prod_{2 \leq t \leq T} p(\gamma_t \mid \gamma_{t-1}, \boldsymbol{\alpha}),\tag{7}$$

where  $p(\mathbf{u}_t \mid \gamma_t)$  is specified by the underlying copula  $c_{\theta_t}(\mathbf{u}_t)$  with parameter  $\theta_t$  specified in (4). More precisely for  $\mathbf{u}_t = (u_{t1}, u_{t2})^\top$  we have:

$$p(\mathbf{u}_t \mid \gamma_t) = c_{\theta_t}(u_{t1}, u_{t2}), \quad \text{where } \theta_t = \tau^{-1}(h(\gamma_t)).\tag{8}$$

*Prior specifications.* For the priors of the copula dependence parameters  $\boldsymbol{\alpha} = (\mu, \phi, \sigma)^\top$  we use standard priors of AR(1) models (see *e.g.* Chib and Greenberg (1994)).

$$\begin{aligned}
(a) \quad & \sigma^2 \sim IG(a, b) \\
(b) \quad & \mu \mid \sigma \sim N(\mu_0, \kappa \sigma^2) \\
(c) \quad & \phi \mid \sigma, \mu \sim U(-1, 1),
\end{aligned} \tag{9}$$

where  $a, b, \mu_0$  and  $\kappa$  are fixed hyper-parameters. Here  $IG(a, b)$  denotes the inverse Gamma distribution with density  $p(\sigma^2) = \frac{b^a}{\Gamma(a)} \left(\frac{1}{\sigma^2}\right)^{a+1} \exp\left(-\frac{b}{\sigma^2}\right)$  and  $U(-1, 1)$  denotes the uniform distribution on  $(-1, 1)$ .

In the next section we show that this prior specification can be used to develop usable expressions for the conditional densities in order to construct a Metropolis-Hastings within Gibbs sampler. The values of the hyper-parameters can be chosen to reflect prior information. A reference prior is available by considering the limiting case when  $a \rightarrow 0$ ,  $b \rightarrow 0$ , and  $\kappa \rightarrow \infty$  giving  $p(\mu, \phi, \sigma^2) \propto \frac{1}{\sigma}$ .

### 3. Posterior Inference

In this section Metropolis-Hastings within Gibbs samplers are developed for the model parameters  $\boldsymbol{\alpha}$  and latent variables  $\boldsymbol{\gamma} = (\gamma_1, \dots, \gamma_T)^\top$  jointly. One sampler consists of naively updating the latent variables  $\boldsymbol{\gamma}$  individually, while a more efficient estimator will be developed using ideas of Liu and Sabatti (2000).

*Updating latent variables  $\boldsymbol{\gamma}$ .* First we introduce some additional notation. Let  $\boldsymbol{\gamma}_A := (\gamma_t : t \in A)$  and  $\boldsymbol{\gamma}_{\setminus A} := \boldsymbol{\gamma}_{A^c}$ . Similarly  $\mathbf{u}_A$  and  $\mathbf{u}_{\setminus A}$  are defined, respectively. From (6) it follows that  $\boldsymbol{\gamma}_A \perp\!\!\!\perp \mathbf{u}_{\setminus A} \mid \mathbf{u}_A, \boldsymbol{\gamma}_{\setminus A}, \boldsymbol{\alpha}$ . Therefore we have

$$\begin{aligned}
p(\boldsymbol{\gamma}_A \mid \mathbf{u}_{1:T}, \boldsymbol{\alpha}, \boldsymbol{\gamma}_{\setminus A}) &= p(\boldsymbol{\gamma}_A \mid \mathbf{u}_A, \boldsymbol{\gamma}_{\setminus A}, \boldsymbol{\alpha}) \\
&\propto p(\mathbf{u}_A, \boldsymbol{\gamma}_A \mid \boldsymbol{\gamma}_{\setminus A}, \boldsymbol{\alpha}) \\
&= p(\boldsymbol{\gamma}_A \mid \boldsymbol{\gamma}_{\setminus A}, \boldsymbol{\alpha}) p(\mathbf{u}_A \mid \boldsymbol{\gamma}_A, \boldsymbol{\gamma}_{\setminus A}, \boldsymbol{\alpha}) \\
&= p(\boldsymbol{\gamma}_A \mid \boldsymbol{\gamma}_{\setminus A}, \boldsymbol{\alpha}) p(\mathbf{u}_A \mid \boldsymbol{\gamma}_A) \quad (\text{by (2) and (3)})
\end{aligned} \tag{10}$$



Thus in a multivariate Gibbs sampler scheme for updating  $\boldsymbol{\gamma}_A$  by using  $p(\boldsymbol{\gamma}_A \mid \mathbf{u}_{1:T}, \boldsymbol{\alpha}, \boldsymbol{\gamma}_{\setminus A})$ , only the observations included in  $\mathbf{u}_A$  are involved. Consequently some computations are simplified.

First a naïve method is described and this sampler is improved by a coarse grid sampler, which takes into account some characteristics of the autoregressive structure for the latent variables.

*Naïve method.* This update of the latent variables  $\boldsymbol{\gamma}$  is based on the *data augmentation principle* introduced in Tanner and Wong (1987) and used for state space models in Carlin, Polson, and Stoffer (1992) or Geweke and Tanizaki (2001). The latent variables are treated as unknown parameters and therefore simulated from the posterior distribution. The full conditional distributions can be derived as follows.

From (10) it follows for  $A = \{t\}$  that

$$p(\gamma_t \mid \mathbf{u}_{1:T}, \boldsymbol{\gamma}_{\setminus t}, \boldsymbol{\alpha}) \propto p(\gamma_t \mid \boldsymbol{\gamma}_{\setminus t}, \boldsymbol{\alpha}) p(\mathbf{u}_t \mid \gamma_t). \quad (11)$$

The Gaussian AR(1) specification for  $\gamma_t$ , given in (5), implies:

$$\gamma_t \mid \boldsymbol{\gamma}_{\setminus t}, \boldsymbol{\alpha} \sim N(\gamma_t^*, \sigma_t^{*2}), \quad (12)$$

where

$$\begin{aligned} \gamma_1^* &:= \mu + \phi(\gamma_2 - \mu); & \sigma_1^{*2} &:= \sigma^2 \\ \gamma_t^* &:= \mu + \frac{\phi}{1 + \phi^2}(\gamma_{t+1} + \gamma_{t-1} - 2\mu); & \sigma_t^{*2} &:= \frac{\sigma^2}{1 + \phi^2}, \text{ for } t = 2, \dots, T-1 \\ \gamma_T^* &:= \mu + \phi(\gamma_{T-1} - \mu); & \sigma_T^{*2} &:= \sigma^2. \end{aligned}$$

Thus, in order to update  $\gamma_t$  given  $\mathbf{u}_{1:T}, \boldsymbol{\gamma}_{\setminus t}, \boldsymbol{\alpha}$ , we perform a Metropolis-Hasting (MH) step for  $\gamma_t$  with proposal density  $N(\gamma_t^*, \sigma_t^{*2})$  and MH-ratio given by:  $\min \left\{ \frac{p(\mathbf{u}_t \mid \tilde{\gamma}_t)}{p(\mathbf{u}_t \mid \gamma_t)}, 1 \right\}$ , where  $\tilde{\gamma}_t$  is the proposal value and  $\gamma_t$  the current value. For more details see the Algorithm 1 in Appendix A.

In spite of its generality and easy implementation, this updating procedure has an important drawback. It shows poor convergence and mixing. One of the potential reason for this could be that the proposal density is far from the posterior one. Improvements will depend on the functional form of the selected copula family. For example Geweke and Tanizaki (2001) suggest different proposal densities for state-space models, which depend on approximations of posterior densities or at least on approximations of the first and

second moments. However, we will follow a different way of improving the mixing of the MCMC algorithm. In the naïve method, the latent variables are updated individually, which does not take into account the autoregressive structure of the latent variables. Therefore we suggest to propagate the updating step into neighbourhoods as a potential improvement.

*Coarse grid method.* The basis of this method can be found in Liu and Sabatti (2000). This generalises the naïve method by taking into account the autoregressive structure of the model. In general this method requires more computational effort, but the behaviour of the simulated Markov chain is improved significantly.

The general idea of the coarse grid method is to find a group of random transformations which leaves distribution such as the posterior or the full conditional distribution invariant. These random transformations are then applied to the naïve sampler, which has then the potential to improve the mixing of the chain. In our case we consider for each  $t = 1, \dots, T$  the following group of transformations:

$$\Gamma_t := \{g_t: g_t(\boldsymbol{\gamma}_{1:T}) = \boldsymbol{\gamma}_{1:T} + \lambda_t \mathbf{b}_t\}, \quad (13)$$

where  $\mathbf{b}_t \in \mathbb{R}^T$  is fixed, and  $\lambda_t \in \mathbb{R}$ . Even though  $\mathbf{b}_t$  can be chosen arbitrary, a convenient choice incorporates some characteristics of the Markovian structure. Motivated by the state-space model example in Liu and Sabatti (2000) we use:

$$\begin{aligned} b_1 &= (1, \frac{\phi}{1+\phi^2}, 0, \dots, 0)^\top \\ b_2 &= (\phi_\ell, 1, \frac{\phi}{1+\phi^2}, 0, \dots, 0)^\top \\ b_t &= (0, \dots, 0, \frac{\phi}{1+\phi^2}, 1, \frac{\phi}{1+\phi^2}, 0, \dots, 0)^\top \quad \text{for } t = 3, \dots, T-2 \\ b_{T-1} &= (0, \dots, 0, \frac{\phi}{1+\phi^2}, 1, \phi_\ell)^\top \\ b_T &= (0, \dots, 0, \frac{\phi}{1+\phi^2}, 1)^\top, \end{aligned} \quad (14)$$

which corresponds to transform  $(\gamma_{t-1}, \gamma_t, \gamma_{t+1})$  into  $(\gamma_{t-1} + \frac{\phi \lambda_t}{1+\phi^2}, \gamma_t + \lambda_t, \gamma_{t+1} + \frac{\phi \lambda_t}{1+\phi^2})$ , for  $t = 3, \dots, T-2$ .

The group of transformations given in (13) and (14) satisfies the condition of Theorem 1 in Liu and Sabatti (2000). Therefore the density  $p(\lambda_t | \mathbf{u}_{1:T}, \boldsymbol{\gamma}_{1:T}, \boldsymbol{\alpha})$  of the random transformation  $\lambda_t$  which leaves the conditional distribution  $p(\boldsymbol{\gamma}_{1:T} | \mathbf{u}_{1:T}, \boldsymbol{\alpha})$  invariant is proportional to

$$p(\lambda_t | \mathbf{u}_{1:T}, \boldsymbol{\gamma}_{1:T}, \boldsymbol{\alpha}) \propto p(g_t(\boldsymbol{\gamma}_{1:T}) | \mathbf{u}_{1:T}, \boldsymbol{\alpha}) \cdot |J_{g_t}(\boldsymbol{\gamma}_{1:T})| \cdot L_t(d\lambda_t), \quad (15)$$

where  $J_{g_t}(\boldsymbol{\gamma}_{1:T})$  is the Jacobian matrix of the transformation and  $L_t$  is the left Haar invariant measure associated to the group of transformations  $\Gamma_t$ . It is easy to verify that  $|J_{\lambda_t}(\boldsymbol{\gamma}_{1:T})| = 1$  and that  $L_t$  is the Lebesgue measure.

Denoting by  $R_T \in \mathbb{R}^{T \times T}$  the matrix with  $(i, j)$ -th element given by  $\phi^{|i-j|}$ ,  $1 \leq i, j \leq T$ , it follows from the AR(1) specification of the latent variables  $\gamma_t$ , in (5) that

$$\boldsymbol{\gamma}_{1:T} | \boldsymbol{\alpha} \sim N_T \left( \mathbf{1}_T \mu, \frac{\sigma^2}{1 - \phi^2} R_T \right), \quad (16)$$

which implies that

$$p(\boldsymbol{\gamma}_{1:T} | \boldsymbol{\alpha}) \propto \exp \left\{ -\frac{1 - \phi^2}{2 \sigma^2} (\boldsymbol{\gamma}_{1:T} - \mathbf{1}_T \mu)^\top R_T^{-1} (\boldsymbol{\gamma}_{1:T} - \mathbf{1}_T \mu) \right\}. \quad (17)$$

From the general model specification, by applying Bayes' Theorem and by simplifying using the equations (2) and (3), we note that

$$\begin{aligned} p(\boldsymbol{\gamma}_{1:T} | \boldsymbol{\alpha}, \mathbf{u}_{1:T}) &\propto p(\mathbf{u}_{1:T} | \boldsymbol{\gamma}_{1:T}, \boldsymbol{\alpha}) p(\boldsymbol{\gamma}_{1:T} | \boldsymbol{\alpha}) \\ &= p(\mathbf{u}_{1:T} | \boldsymbol{\gamma}_{1:T}) p(\boldsymbol{\gamma}_{1:T} | \boldsymbol{\alpha}). \end{aligned}$$

This yields

$$\begin{aligned} p(g_t(\boldsymbol{\gamma}_{1:T}) | \boldsymbol{\alpha}, \mathbf{u}_{1:T}) &\propto \exp \left\{ -\frac{1 - \phi^2}{2 \sigma^2} (\boldsymbol{\gamma}_{1:T} + \lambda_t \mathbf{b}_t - \mathbf{1}_T \mu)^\top R_T^{-1} (\boldsymbol{\gamma}_{1:T} + \lambda_t \mathbf{b}_t - \mathbf{1}_T \mu) \right\} \\ &\quad \cdot p(\mathbf{u}_{1:T} | g_t(\boldsymbol{\gamma}_{1:T})). \quad (18) \end{aligned}$$

Using (15) it follows that

$$\begin{aligned}
p(\lambda_t \mid \boldsymbol{\gamma}_{1:T}, \mathbf{u}_{1:T}, \boldsymbol{\alpha}_\ell) &\propto \exp \left\{ -\frac{1-\phi^2}{2\sigma^2} (\boldsymbol{\gamma}_{1:T} + \lambda_t \mathbf{b}_t - \mathbf{1}_T \boldsymbol{\mu})^\top R_T^{-1} (\boldsymbol{\gamma}_{1:T} + \lambda_t \mathbf{b}_t - \mathbf{1}_T \boldsymbol{\mu}) \right\} \\
&\quad \cdot p(\mathbf{u}_{1:T} \mid g_t(\boldsymbol{\gamma}_{1:T})) \\
&= \exp \left\{ -\frac{1-\phi^2}{2\sigma^2} (\lambda_t^2 \mathbf{b}_t^\top R_T^{-1} \mathbf{b}_t + 2 (\boldsymbol{\gamma}_{1:T} - \mathbf{1}_T \boldsymbol{\mu})^\top R_T^{-1} \mathbf{b}_t \lambda_t) \right\} \\
&\quad \cdot p(\mathbf{u}_{1:T} \mid g_t(\boldsymbol{\gamma}_{1:T})) \\
&\propto \exp \left\{ -\frac{(1-\phi^2) \mathbf{b}_t^\top R_T^{-1} \mathbf{b}_t}{2\sigma^2} \left( \lambda_t + \frac{(\boldsymbol{\gamma}_{1:T} - \mathbf{1}_T \boldsymbol{\mu})^\top R_T^{-1} \mathbf{b}_t}{\mathbf{b}_t^\top R_T^{-1} \mathbf{b}_t} \right)^2 \right\} \\
&\quad \cdot p(\mathbf{u}_{1:T} \mid g_t(\boldsymbol{\gamma}_{1:T})) \\
&= \exp \left\{ -\frac{1}{2V_t^{*2}} (\lambda_t - \lambda_t^*)^2 \right\} \cdot p(\mathbf{u}_{1:T} \mid g_t(\boldsymbol{\gamma}_{1:T})) \tag{19}
\end{aligned}$$

with

$$\lambda_t^* := -\frac{(\boldsymbol{\gamma}_{1:T} - \mathbf{1}_T \boldsymbol{\mu})^\top R_T^{-1} \mathbf{b}_t}{\mathbf{b}_t^\top R_T^{-1} \mathbf{b}_t} \text{ and } V_t^{*2} := \frac{\sigma^2}{(1-\phi^2) \mathbf{b}_t^\top R_T^{-1} \mathbf{b}_t}. \tag{20}$$

Now, by applying equation (19), we can show that, for all  $a \in \mathbb{R}$   $p(\lambda_t \mid \boldsymbol{\gamma}_{1:T}, \mathbf{u}_{1:T}, \boldsymbol{\alpha}) = p(\lambda_t + a \mid \boldsymbol{\gamma}_{1:T} - a \mathbf{1}_T, \mathbf{u}_{1:T}, \boldsymbol{\alpha})$ . holds. This corresponds to the transformation-invariant property (1) of Liu and Sabatti (2000). Therefore we can apply Theorem 2 of Liu and Sabatti (2000) and use MH steps for updating the latent variables. Therefore, we choose  $\lambda_t$  deriving from a MH step with proposal distribution  $N(\lambda_t^*, V_t^{*2})$  and MH-ratio

$$\min \left\{ \frac{p(\mathbf{u}_{A_t} \mid \boldsymbol{\gamma}_{A_t} + \lambda_t \mathbf{b}_{A_t})}{p(\mathbf{u}_{A_t} \mid \boldsymbol{\gamma}_{A_t})}, 1 \right\}, \tag{21}$$

where  $A_t = \{j \in \{1, \dots, T\} : b_{t,j} \neq 0\}$  and  $b_{t,j}$  is the  $j$ th component of  $\mathbf{b}_t$ .

In summary, the complete coarse grid update proceeds as follows:

- For current values of  $\boldsymbol{\alpha}$  update  $\boldsymbol{\gamma}_{1:T}$  using the naïve method.
- For  $t = 1, \dots, T$ , simulate a  $\lambda_t$  by using an MH step as described, compute  $g_t(\boldsymbol{\gamma}_{1:T}) = \boldsymbol{\gamma}_{1:T} + \lambda_t \mathbf{b}_t$  and use them as new values for  $\boldsymbol{\gamma}_{1:T}$ .

Note that the MH-ratio in (21) can be computed quickly due to the large number of zeroes in  $\mathbf{b}_t$ . More details are given in the Algorithm 2 in Appendix A. In Appendix B we give simple equations for computing  $\lambda_t^*$  and  $V_t^{*2}$ .

*Updating model parameters  $\alpha$  of the latent variables.* The conditional independence condition (6b) implies that once the latent variables are updated, the new values for the parameters  $\mu, \phi, \sigma$  depend only on the current values of the latent variables  $\gamma_{1:T}$  and not on the observed values  $\mathbf{u}_{1:T}$ , *i.e.*

$$p(\mu, \phi, \sigma \mid \gamma_{1:T}, \mathbf{u}_{1:T}) = p(\mu, \phi, \sigma \mid \gamma_{1:T}). \quad (22)$$

Now, the appropriate conditional distributions for the Gibbs sampler are as follows:

For  $\mu$  and  $\sigma$ : from (22) and the prior specification (9),

$$\begin{aligned} p(\mu, \sigma^2 \mid \mathbf{u}_{1:T}, \gamma_{1:T}, \phi) &= p(\mu, \sigma^2 \mid \gamma_{1:T}, \phi) \\ &\propto p(\gamma_{1:T} \mid \mu, \phi, \sigma) p(\mu, \sigma) \\ &\propto p(\mu, \sigma^2) \left( \frac{1}{\sigma^2} \right)^{\frac{T}{2}} \exp \left\{ -\frac{A}{2\sigma^2} \left( \mu - \frac{m_1}{A} \right)^2 \right\} \\ &\quad \cdot \exp \left\{ -\frac{1}{2\sigma^2} \left( m_2 - \frac{m_1^2}{A} \right) \right\}, \end{aligned}$$

where  $A := 1 - \phi^2 + (T-1)(1-\phi)^2$ ,  $m_1 := (1-\phi^2)\gamma_t + (1-\phi) \sum_{t=2}^T (\gamma_{\ell,t} - \phi\gamma_{t-1})$

and  $m_2 := (1-\phi^2)\gamma_t^2 + \sum_{t=2}^T (\gamma_t - \phi\gamma_{t-1})^2$ . Therefore integrating over  $\mu$ , and using the prior specification (9) results in

$$p(\sigma^2 \mid \gamma_{1:T}, \phi) \propto \left( \frac{1}{\sigma^2} \right)^{\frac{T-1}{2} + a_\ell + 1} \exp \left\{ -\frac{B+b}{\sigma^2} \right\},$$

with  $B := \frac{1}{2} \left( m_2 - \frac{m_1^2}{A} \right)$ . Here  $a$  and  $b$  are the hyper-parameters of the prior specification. This can be further identified as

$$\sigma^2 \mid \gamma_{1:T}, \phi \sim IG \left( \frac{T-1}{2} + a, B+b \right). \quad (23)$$

Now, for the conditional distribution of  $\mu$ , note that

$$p(\mu \mid \gamma_{1:T}, \phi, \sigma^2) \propto p(\mu \mid \sigma^2) \exp \left\{ -\frac{A}{2\sigma^2} \left( \mu - \frac{m_1}{A} \right)^2 \right\} \quad (24)$$

by using the prior specification (9). Easy calculation shows that

$$\mu \mid \gamma_{1:T}, \phi, \sigma^2 \sim N(\tilde{\mu}^*, \tilde{\sigma}_\mu^{*2}), \quad (25)$$

where  $\tilde{\mu}^* = \frac{\kappa m_1 + \mu^*}{1 + \kappa A}$  and  $\tilde{\sigma}_\mu^{*2} = \sigma^2 \left( \frac{\kappa}{1 + \kappa A} \right)$ .

If  $a \rightarrow 0$ ,  $b \rightarrow 0$ , and  $\kappa \rightarrow \infty$ , *i.e.* when non informative priors are used, the conditional distribution simplifies to

$$\sigma^2 \mid \gamma_{1:T}, \phi \sim IG\left(\frac{T-1}{2}, B\right) \quad \text{and} \quad \mu \mid \gamma_{1:T}, \phi, \sigma^2 \sim N\left(\frac{m_1}{A}, \frac{\sigma^2}{A}\right).$$

For  $\phi$  we have that

$$p(\phi \mid \gamma_{1:T}, \mu_\ell) \propto \varphi_{\mu, \frac{1}{1-\phi^2}}(\gamma_1) \cdot \varphi_{\frac{S_1}{S_0}, \frac{1}{S_0}}(\phi) \cdot \mathbf{1}_{[-1,1]}(\phi), \quad (26)$$

where  $S_0 = \sum_{t=2}^T (\gamma_t - \mu)^2$ , and  $S_1 = \sum_{t=2}^T (\gamma_t - \mu)(\gamma_{t-1} - \mu)$ . Here  $\varphi_{\mu, \sigma^2}$  denotes the  $N(\mu, \sigma^2)$  density.

Therefore, with the specified prior, *i.e.*  $p(\phi) \propto 1$ , we use an MH-step with a  $N\left(\frac{S_1}{S_0}, \frac{1}{S_0}\right)$  distribution truncated to  $[-1, 1]$  as proposal distribution. Since the proposal distribution is close to the posterior one, the acceptance rate is high. Moreover the acceptance rates approach one when the sample size grows since correction is only needed for the initial state. Algorithm 3 in Appendix A gives a summary for updating the latent model parameters.

#### 4. Simulation Study

We investigate in the following the behaviour of the two MCMC samplers developed in Section 3 for a Gaussian, double Clayton and double Gumbel copula with different stochastic time-varying copula dependency.

*Gaussian copula.* A first set of scenarios for the simulation involves Gaussian SCAR copula models. The sample size was fixed to  $T = 1000$ , and the simulated Markov chains consist of 100 000 iterations using the naïve (N) and coarse grid (CG) sampler, respectively. For a graphical check of the behaviour of the generated Markov chain using the CG sampler, some plots are presented for a single simulated data set in Figure 2.

—Figure 2 about here—

Figure 2 illustrates the convergence of the CG sampler for the copula dependence parameters  $\alpha$ . The true values are reasonably close to the maximum a posteriori (MAP) estimates. The MCMC traces seem stationary and the estimated autocorrelation function decreases quickly for  $\mu$  and  $\phi$  and reasonably for  $\sigma$ . For the posterior analysis we take every 50th iterations to have low autocorrelations for all parameters.

Additionally, since the latent variables are treated as unknown parameters, credibility intervals can be computed for them. For many parametric copula families, an analytic one-to-one relationship between Kendall's  $\tau$  and the copula parameter is known. Therefore we can construct credible intervals for the time varying Kendall's  $\tau_t$ . The 90% point-wise credible intervals for  $\tau_t$  for the last 200 observations for a simulated data set are presented in Figure 3, demonstrating that the true  $\tau_t$  values (dotted line) is well covered. Further we see wider credible intervals for lower values of  $\tau_t$  compared to higher values. This is due to the fact that an observation  $(u_{1t}, u_{2t})$  where  $u_{1t} \approx u_{2t}$  can be associated with high dependence but also with low dependence, while an observation where  $u_{1t} \ll u_{2t}$  or  $u_{1t} \gg u_{2t}$  can only be associated with lower dependence. Thus low values of the Kendall's  $\tau$ , i.e. lower dependence, are less precisely estimated. Additionally we see that the credible intervals are non symmetric, reflecting nonsymmetric posterior distributions of  $\tau_t$ .

—Figure 3 about here—

In order to check the validity of the naïve (N) and the coarse grid (CG) sampler, as well as the improved efficiency of the CG sampler over the N sampler, we simulate 100 data sets for several parameter configurations. A summary of the behaviour for the two samplers is presented in Table 1.

—Table 1 about here—

In addition to estimated posterior quantiles, posterior mean and mode, averaged over the 100 simulation we also report the average *effective sample size* (EFS) for the two samplers and its ratio. The EFS for R MCMC iterations is the number of i.i.d. samples necessary to achieve the accuracy achieved by the R MCMC iterations. More precisely the variance of the

posterior mean estimate  $\bar{\theta}$  for the parameter  $\theta$  by using  $R$  MCMC iterations  $\theta^{(r)}$  is given by:

$$\text{Var}(\bar{\theta}) = \frac{\text{Var}(\theta^{(r)})}{R} + \frac{1}{R^2} \sum_{r \neq s} \text{Cov}(\theta^{(r)}, \theta^{(s)}). \quad (27)$$

The first term on the right side of equation (27) can be interpreted as the variance of  $\bar{\theta}$  if the sample is i.i.d. Therefore, the effective sample size is defined such that

$$EFS := \frac{\text{Var}(\theta^{(r)})}{\text{Var}(\bar{\theta})} = \frac{R}{1 + \frac{1}{R} \sum_{r \neq s} \text{Cor}(\theta^{(r)}, \theta^{(s)})}.$$

From the posterior summaries we see that both samplers are estimating the true parameters quite well over all parameter combinations. As expected the effective sample size is quite low for  $\phi$  and  $\sigma$ . However the CG sampler is always more efficient than the N sampler. The improvement is larger for strong dependencies (high  $\phi$ ) than for small dependencies (low  $\phi$ ). In the cases of large  $\phi$  an improvement of about 40% is observed.

*Clayton copula.* A second example involves a mixture of Clayton copulas, hereafter called double Clayton. The definition of a double Clayton is as follows: Denoting by  $C(\theta)$  a Clayton copula with parameter  $\theta \in [0, \infty)$ , A random vector  $(U, V) \in [0, 1]^2$  is distributed as a double Clayton  $DC(\theta)$  with  $\theta \in \mathbb{R}$  if  $(U, V)$  is distributed as a Clayton  $C(\theta)$  for  $\theta \geq 0$  and  $(U, 1 - V)$  is distributed as a Clayton  $C(-\theta)$  for  $\theta < 0$ . This copula can also be viewed as a rotated copula by 90 degrees. This idea can be applied to other copula families such as the Gumbel copula and further examples can be found in Joe (1997).

The motivation for using these copulas rather than a simple Clayton or Gumbel is to avoid problems when negative dependence occurs. This situation could happen if  $\phi$  is negatively estimated for some MCMC iterations. From a practical point of view, if negative dependence is not present in the data, we will expect at most a few negative  $\phi$  iterations. The advantage of this formulation is that it permits us to use the same sampling algorithms. Therefore we can compare for example a Gaussian SCAR copula model to a Clayton or Gumbel ones without having to consider the sign of the dependence. Figure 4 gives density contours for a bivariate double Clayton



distribution with standard normal margins (top row) and a double Gumbel distribution with standard normal margins (bottom row), respectively.

—Figure 4 about here—

We now present a similar table of the simulation results for the posterior estimation for double Clayton SCAR copula models as in the Gaussian case.

—Table 2 about here—

Table 2 shows similar good behaviour of the posterior mean and MAP estimates for the both samplers. The EFS are similar than in the Gaussian SCAR case.

*Gumbel Example.* For the same setup of the parameters as in the Gaussian and in the Clayton examples, the behaviour is similar (not shown). Further, we want to here explore a different scenario where the Kendall's  $\tau_t$  varies with a high frequency; therefore, we propose scenarios where the autocorrelation parameter ( $\phi$ ) is negative. The results of this simulation setup of double Gumbel SCAR models are in presented in Table 3.

—Table 3 about here—

It shows good behaviour of the posterior mean and MAP estimates for the both samplers. The EFS are lower than in the Gaussian and in the Clayton SCAR cases and the improvement of the CG sampler over the N sampler is not so pronounced, but still visible with improvements of 18%-57% for almost all parameters.

Overall we showed that both samplers are accurate and stable for Gaussian, double Clayton and double Gumbel SCAR copulas for different parameter configurations. There is an efficiency gain for using the CG sampler over the N sampler. This gain is higher for low frequency Kendall's  $\tau$  variation *i.e.* the autocorrelation ( $\phi$ ) is positive.

We like to mention that a fast implementation is available. More precisely, 100 000 iteration for a bivariate time series of size 1 00 takes about 15 minutes on a Intel Core Duo T7500 running in 2.2GHz, and using only one processor.

## 5. Empirical Applications

In order to see how the proposed estimator works with a real data set, we consider two bivariate financial time series. The first one is the data set used in Engle (2002), *i.e.* the log returns of the Industrial Dow Jones (DJI) and the NASDAQ from March 23th, 1990 to March 26th, 2000. The second data set comes from the Standard & Poors 500 (S&P500) and the Deutscher Aktien Index (DAX) from June 6th, 2002 until November 6th, 2009.

First we estimate two separate univariate GARCH(1,1) models with Gaussian innovations (see *e.g.* Bollerslev (1986)) for each margin to account for marginal time dependence. A summary of the marginal estimates is reported in Table 4. The standardized residuals behave like an i.i.d. standard normal sample and as such, we believe marginal time dependence has been sufficiently removed. Copula data was derived using the probability integral transformation. We leave out the two last years in order to evaluate the forecasting performance of several SCAR copula models using the CG sampler for prediction.

—Table 4 about here—

The estimated posterior quantiles, mean and mode for the time varying copula parameters using Gaussian, Clayton and Gumbel SCAR models are presented in Table 5. Here a CG sampler was run using 100 000 MCMC iterations, we found a burnin of 5 000 iterations and a subsampling of every 50th iteration sufficient, yielding 1 900 recorded MCMC iterations.

—Table 5 about here—

We see that the persistence of the “dependence shocks” (size of  $\phi$ ) varies with the SCAR copula model as well as for the data set considered. For the DOW Jones-NASDAQ data the Clayton (Gaussian) SCAR copula model has the lowest (highest)  $\phi$ , while for the S&P500-DAX data the Clayton (Gumbel) SCAR copula model has the lowest (highest) one. From the estimation of  $\mu$ , we observe higher mean time dependence between DOW Jones NASDAQ data compared with S&P500-DAX data. This might be explained by the fact that both DOW Jones and NASDAQ are indices from the same stock exchange market, namely the New York Stock Exchange, while this is not the case for the S&P500 and DAX indices. As noted we observe that the

“dependence shocks” decay slower in DOW Jones-NASDAQ data compared to the S&P500 - DAX data. This difference might follow again from the markets where these indices are traded.

To assess which of the SCAR copula models is the most appropriate one we compare sum of squared distances between the corresponding SCAR Kendall’s  $\tau_t$  estimates and rolling window Kendall’s  $\tau_t$  estimates. Here the rolling window estimates serve as empirical estimates and therefore we select the SCAR model which provides the lowest sum of squared distances. The corresponding sum of squared distances are given in Table 6 with the lowest values per window size in bold numbers. We see that for the Dow Jones-NASDAQ data the Gaussian SCAR model give the lowest for two window sizes and is of lowest equal magnitude for a window size of 250. Using this measure we would therefore choose a Gaussian SCAR copula model for this data set. For the S&P500-DAX data the Gumbel SCAR model would be chosen for a rolling window size of 50, while a Clayton SCAR model would be more appropriate if one considers a window size of 125 or 250. Here we make the choice of a window size of 50, since this does not provide oversmoothing of the copula time dependence measure.

—Table 6 about here—

To assess in more detail the fit of the different SCAR specifications we now compare the posterior mean of Kendall’s  $\tau_t$  with the corresponding rolling window estimates using window size 50 (see top row of Figure 6). This confirms our model choice of a Gaussian SCAR model for the Dow Jones-NASDAQ data and a Gumbel SCAR model for the S&P500-DAX data. In the bottom panels of Figure 6 we show Kendall’s  $\tau_t$  estimates for the chosen models together with point wise 90% credible intervals and the rolling window estimate with window size 50. They show that the posterior mean estimates are close to the rolling window estimate. Further non overlapping credible interval estimates at different time points show the need for time varying copula dependence.

—Figure 6 about here—

Finally we consider the predictive capability of the chosen SCAR models by forecasting Kendall’s  $\tau_t$ . For this we simulate for the  $r$ th MCMC SCAR

model parameter vector value corresponding latent variables  $\gamma_t^{(r)}$  for  $t \geq T$ , which we then transform to the corresponding Kendall's  $\tau_t$  values. We then use these values to construct corresponding posterior mean estimates together with point-wise 90% credible estimates. The results are shown in Figure 7. Both panels indicate the high predictive capability of the chosen SCAR models for the two data sets, since the point-wise predictive intervals cover the rolling window estimate

—Figure 7 about here—

## 6. Summary and outlook

In this paper we considered SCAR copula models, which are copula models allowing for stochastic time-varying dependence driven by time dependent latent variables. The latent variables follow a stationary AR(1) model and the inverse Fisher transform relates the latent variable at time  $t$  to the Kendall's  $\tau$  at time  $t$ . Any bivariate copulas can be taken. For the simulation and application we chose the Gaussian, Clayton and Gumbel copula.

We took a Bayesian approach using MCMC sampling for estimation and inference, since maximum likelihood is not tractable. The reason is that the integration over  $T$  latent variables is computationally infeasible. In addition the Bayesian approach allows easy construction of credible intervals, enabling the assessment of the precision of point estimates. Further using the MCMC iterates we can easily construct point-wise credible intervals for interesting time-varying quantities such as Kendall's  $\tau_t$  or the tail dependence coefficient  $\lambda_t$ , as long there are simple relationships between the copula parameter  $\theta_t$ , Kendall's  $\tau_t$  and the tail dependence coefficient  $\lambda_t$ . This was the case for the copula families considered.

The need of having to update all latent variables requires special care when a MCMC algorithm is designed. Here we developed two sampling schemes. One is a naïve sampling scheme for the latent variables which uses individual Metropolis-Hasting steps. This sampler however exhibit large autocorrelations. We improved this sampling scheme by developing a coarse grid sampler. Here we found a random transformation which satisfied the conditions of Liu and Sabatti (2000) and is easy to sample. According to Liu and Sabatti (2000) the application of this random transformation does not change the needed full conditional. In a simulation study we showed that

this procedure improves the mixing of the MCMC chain as measured by the effective sample size.

In two applications to financial stock indices we demonstrated the need to incorporate time varying dependence. We applied the coarse grid sampler for SCAR models with Gaussian, Clayton and Gumbel copulas. To assess the fit of these different SCAR models we compared the posterior mean of Kendall's  $\tau_t$  to empirical estimates of  $\tau_t$ . The empirical estimates of  $\tau_t$  were computed using a simple rolling window with fixed window size. Here we like to note that the SCAR model formulation allows for easy comparison of different copula families, since the time dependence is modelled in terms of Kendall's  $\tau_t$  and not with respect to the parameter  $\theta_t$  which specifies the copula. Finally we investigated the capabilities of the chosen SCAR model with respect to predicting future Kendall's  $\tau_t$ .

There are several interesting open problems to investigate. First one can develop a joint Bayesian analysis of marginal models such as GARCH or stochastic volatility (SV) models as suggested in Ausin and Lopes (2009) using a nonstochastic dependence dynamics. At the moment the samplers have been formulated for bivariate time series, however we can extend this by using multivariate copulas. A very flexible class of such models has been suggested and applied to financial time series data by Aas, Czado, Frigessi, and Bakken (2009). It uses the pair-copula construction (PCC) method and a recent survey of these PCC models are given in Czado (2010). The advantage of such an approach is that the model is formulated with respect to conditional parameters, which in contrast to correlation parameters can be chosen independently. Time varying dependence can be incorporated in a similar fashion as for the SCAR copula models. Efficient Bayesian inference for such models are currently being studied. Finally more suitable model comparison criteria for non nested SCAR models are needed. One possible approach here would be to investigate Bayesian adaptations to non nested model comparison tests such as suggested by Vuong (1989) and Clarke (2007). A first such adaptation was provided in Czado, Schabenberger, and Erhardt (2009).

#### *Acknowledgements*

Claudia Czado and Carlos Almeida gratefully acknowledge the financial support from the Deutsche Forschungsgemeinschaft (Cz 86/1-3: Statistical inference for high dimensional dependence models using pair-copulas).

## References

- Aas, K., Czado, C., Frigessi, A., Bakken, H., 2009. Pair-copula constructions of multiple dependence. *Insurance: Mathematics and Economics* 44 (2), 182 – 198.
- Ausin, M. C., Lopes, H. F., 2009. Time-varying joint distribution through copulas. *Computational Statistics & Data Analysis* In Press.
- Bauwens, L., Laurent, S., Rombouts, J. V. K., 2006. Multivariate GARCH models: a survey. *Journal of Applied Econometrics* 21 (1), 79–109.
- Bollerslev, T., 1986. Generalized autoregressive conditional heteroskedasticity. *J. Econometrics* 31 (3), 307–327.
- Carlin, B. P., Polson, N. G., Stoffer, D. S., 1992. A Monte Carlo approach to nonnormal and nonlinear state-space modeling. *Journal of the American Statistical Association* 87 (418), 493–500.
- Cherubini, U., Luciano, E., Vecchiato, W., 2004. Copula methods in finance. Wiley Finance Series. John Wiley & Sons Ltd., Chichester.
- Chib, S., Greenberg, E., 1994. Bayes inference in regression models with ARMA(p, q) errors. *Journal of Econometrics* 64 (1-2), 183 – 206.
- Clarke, K., 2007. A simple distribution-free test for nonnested hypotheses. *Political Analysis* 15 (3).
- Czado, C., 2010. Pair-copula constructions of multivariate copulas. In: Durante, F., Härdle, W., Jaworki, P., Rychlik, T. (Eds.), *Workshop on Copula Theory and its Applications*. Springer, Dordrecht.
- Czado, C., Schabenberger, H., Erhardt, V., 2009. Nonnested model selection for spatial count regression models with application to health insurance. To appear in *Statistical Papers*.
- Dias, A., Embrechts, P., 2004. Change-point analysis for dependence structures in finance and insurance. In: Szegoe, G. (Ed.), *Risk Measures for the 21st Century*. Wiley Finance Series, pp. 32–335.

- Engle, R., 2002. Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models. *Journal of Business & Economic Statistics* 20 (3), 339–350.
- Erb, C. B., Harvey, C. R., Viskanta, T. E., 1994. Forecasting international equity correlations. *Financial Analysts Journal* 50 (6), 32 – 45.
- Florens, J.-P., Mouchart, M., Rolin, J.-M., 1990. Elements of Bayesian statistics. Vol. 134 of Monographs and Textbooks in Pure and Applied Mathematics. Marcel Dekker Inc., New York.
- Genest, C., Favre, A.-C., 2007. Everything you always wanted to know about copula modeling but were afraid to ask. *Journal of Hydrological Engineering* 12 (4), 347–368.
- Genest, C., Ghoudi, K., Rivest, L.-P., 1995. A semiparametric estimation procedure of dependence parameters in multivariate families of distributions. *Biometrika* 82 (3), 543–552.
- Geweke, J., Tanizaki, H., 2001. Bayesian estimation of state-space models using the metropolis-hastings algorithm within gibbs sampling. *Computational Statistics & Data Analysis* 37 (2), 151 – 170.
- Giacomini, E., Härdle, W., Spokoiny, V., 2009. Inhomogeneous dependence modeling with time-varying copulae. *J. Bus. Econom. Statist.* 27 (2), 224–234.
- Hafner, C., Manner, H., 2008. Dynamic stochastic copula models: Estimation, inference and applications, METEOR Research Memorandum (RM) 07/043.
- Harvey, A., Ruiz, E., Shephard, N., 1994. Multivariate stochastic variance models. *The Review of Economic Studies* 61 (2), 247–264.
- Joe, H., 1996. Families of  $m$ -variate distributions with given margins and  $m(m-1)/2$  bivariate dependence parameters. In: *Distributions with fixed marginals and related topics* (Seattle, WA, 1993). Vol. 28 of IMS Lecture Notes Monogr. Ser. Inst. Math. Statist., Hayward, CA, pp. 120–141.
- Joe, H., 1997. Multivariate models and dependence concepts. Vol. 73 of *Monographs on Statistics and Applied Probability*. Chapman & Hall, London.

- Joe, H., 2005. Asymptotic efficiency of the two-stage estimation method for copula-based models. *J. Multivariate Anal.* 94 (2), 401–419.
- Jondeau, E., Rockinger, M., 2006. The Copula-GARCH model of conditional dependencies: An international stock market application. *Journal of International Money and Finance* 25, 827 – 853.
- Kim, G., Silvapulle, M. J., Silvapulle, P., 2007. Comparison of semiparametric and parametric methods for estimating copulas. *Computational Statistics & Data Analysis* 51 (6), 2836 – 2850.
- Liesenfeld, R., Richard, J.-F., 2003. Estimation of dynamic bivariate mixture models: comments on Watanabe (2000). *J. Bus. Econom. Statist.* 21 (4), 570–576.
- Liu, J. S., Sabatti, C., 2000. Generalised Gibbs sampler and multigrid Monte Carlo for Bayesian computation. *Biometrika* 87 (2), 353–369.
- Liu, Y., Luger, R., 2009. Efficient estimation of copula-GARCH models. *Computational Statistics & Data Analysis* 53 (6), 2284 – 2297, the Fourth Special Issue on Computational Econometrics.
- Longin, F., Solnik, B., 1995. Is the correlation in international equity returns constant: 1960-1990? *Journal of International Money and Finance* 14 (1), 3 – 26.
- Manner, H., Candelon, B., 2007. Testing for asset market linkages: A new approach based on time-varying copulas, METEOR Research Memorandum (RM) 07/052.
- Manner, H., Reznikova, O., 2009. Time-varying copulas: a survey, Institut de Statistique UCL, DP0917.
- Mercurio, D., Spokoiny, V., 2004. Statistical inference for time-inhomogeneous volatility models. *Ann. Statist.* 32 (2), 577–602.
- Nelsen, R. B., 2006. An introduction to copulas, 2nd Edition. Springer Series in Statistics. Springer, New York.
- Patton, A. J., 2006. Estimation of multivariate models for time series of possibly different lengths. *Journal of Applied Econometrics* 21 (2), 147–173.



- Richard, J.-F., Zhang, W., 2007. Efficient high-dimensional importance sampling. *J. Econometrics* 141 (2), 1385–1411.
- Sklar, M., 1959. Fonctions de répartition à  $n$  dimensions et leurs marges. *Publ. Inst. Statist. Univ. Paris* 8, 229–231.
- Tanner, M. A., Wong, W. H., 1987. The calculation of posterior distributions by data augmentation. *Journal of the American Statistical Association* 82 (398), 528–540.
- Taylor, S. J., 1986. *Modelling Financial Time Series*. JohnWiley and Sons.
- Tse, Y. K., Tsui, A. K. C., 2002. A multivariate generalized autoregressive conditional heteroscedasticity model with time-varying correlations. *J. Bus. Econom. Statist.* 20 (3), 351–362.
- Vuong, Q. H., 1989. Likelihood ratio tests for model selection and nonnested hypotheses. *Econometrica* 57 (2), 307–333.
- Yu, J., Meyer, R., 2006. Multivariate stochastic volatility models: Bayesian estimation and model comparison. *Econometric Reviews* 25, 361–384.

## Appendix A. Algorithms

---

### Algorithm 1 Updating latent variables (naïve method)

---

**Input:**  $\gamma_t^{(r-1)}, 1 \leq t \leq T$

**Output:**  $\gamma_t^{(r)}, 1 \leq t \leq T$

- 1: **for**  $t \leftarrow 1, T$  **do**
  - 2:     Simulate  $\tilde{\gamma}_t$  from  $N(\gamma_t^*, \sigma_t^{*2})$  ▷ See Eq. (12)
  - 3:     Compute  $k = \min \left\{ \frac{p(\mathbf{u}_t | \tilde{\gamma}_t)}{p(\mathbf{u}_t | \gamma_t^{(r-1)})}, 1 \right\}$
  - 4:     MH-Step:
  - 5:     **if**  $\text{unif}(\mathbf{0}, \mathbf{1}) \leq k$ ,  $\gamma_t^{(r)} \leftarrow \tilde{\gamma}_t$
  - 6:     **else**  $\gamma_t^{(r)} \leftarrow \gamma_t^{(r-1)}$
  - 7: **end for**
- 

---

### Algorithm 2 Updating latent variables (coarse grid method)

---

**Input:**  $\gamma_t^{(r-1)}, 1 \leq t \leq T$

**Output:**  $\gamma_t^{(r)}, 1 \leq t \leq T$

- 1: **for**  $t \leftarrow 1, T$  **do**
  - 2:     Simulate  $\tilde{\lambda}_t$  from  $N(\lambda_t^*, V_t)^*$  ▷ See Eq. (20)
  - 3:     Set  $\tilde{\gamma}_{1:T} = \gamma_{1:T}^{(r-1)} + \tilde{\lambda}_t \mathbf{b}_t$
  - 4:     Compute  $k = \min \left\{ \frac{p(\mathbf{u}_{1:T} | \tilde{\gamma}_{1:T})}{p(\mathbf{u}_{1:T} | \gamma_{1:T}^{(r-1)})}, 1 \right\}$
  - 5:     MH-Step:
  - 6:     **if**  $\text{unif}(\mathbf{0}, \mathbf{1}) \leq k$ ,  $\gamma_{1:T}^{(r)} \leftarrow \tilde{\gamma}_{1:T}$
  - 7:     **else**  $\gamma_{1:T}^{(r)} \leftarrow \gamma_{1:T}^{(r-1)}$
  - 8: **end for**
-

---

**Algorithm 3** Updating parameters

---

**Input:**  $\mu^{(r-1)}, \phi^{(r-1)}, \sigma^{(r-1)}$ **Output:**  $\mu^{(r)}, \phi^{(r)}, \sigma^{(r)}$ 1: **Update** of  $\mu$ :2:  $\sigma^{2(r)}$  from  $IG(\frac{T-1}{2} + a, B + b)$  ▷ see Eq. (23)3: **Update** of  $\mu$ :4:  $\mu^{(r)}$  from  $N(\tilde{\mu}^*, \tilde{\sigma}_\mu^{*2})$  ▷ see Eq. (25)5: **Update** of  $\phi$ 6: Simulate  $\tilde{\phi}$  from  $N\left(\frac{S_1}{S_0}, \frac{1}{S_0}\right) \cdot \mathbf{1}_{[-1,1]}(\tilde{\phi})$  ▷ See Eq. (26)7: Compute  $k = \min \left\{ \frac{\varphi_{\mu, \frac{1}{1-\phi^2}}(\gamma_1^{(r)})}{\varphi_{\mu, \frac{1}{1-\phi^{(r-1)}^2}}(\gamma_1^{(r)})}, 1 \right\}$ 8: **if**  $\text{unif}(0,1) \leq k$ ,  $\phi^{(r)} \leftarrow \tilde{\phi}$ 9: **else**  $\phi^{(r)} \leftarrow \phi^{(r-1)}$ 

---

**Appendix B. Conditional expectations and variances for the coarse grid method**

The values of  $\lambda_t^*$  and  $V_t^*$  for the definition of  $\mathbf{b}_t$  as in equation (14).

$$\text{For } t = 1, 2: \quad \begin{aligned} \lambda_t^* &= -(\gamma_t - \mu) + \phi^2(\gamma_{t+2} - \mu) \\ V_t^* &= \sigma^2(1 + \phi^2) \end{aligned}$$

$$\text{For } t = 3, \dots, T-2: \quad \begin{aligned} \lambda_t^* &= -(\gamma_t - \mu) + \frac{\phi^2(\gamma_{t-2} + \gamma_{t+2} - 2\mu)}{1 + \phi^4} \\ V_t^* &= \frac{\sigma^2(1 + \phi^2)}{1 + \phi^4} \end{aligned}$$

$$\text{For } t = T-1, T: \quad \begin{aligned} \lambda_t^* &= -(\gamma_t - \mu) + \phi^2(\gamma_{t-2} - \mu) \\ V_t^* &= \sigma^2(1 + \phi^2) \end{aligned}$$

CG sampler								
	True	2.5%	50%	97.5%	mean	mode	EFS	$\frac{\text{EFS}_{cg}}{\text{EFS}_N}$
$\mu$	1.0	0.9339	1.0050	1.0736	1.0049	1.0053	31954.33	3.2498
$\phi$	0.9	0.8031	0.8932	0.9491	0.8890	0.8994	1214.65	1.3400
$\sigma$	0.1	0.0754	0.1064	0.1466	0.1077	0.1044	750.73	1.5327
$\mu$	1.0	0.9657	1.0013	1.0367	1.0013	1.0011	4582.14	1.3387
$\phi$	0.1	-0.3004	0.0855	0.4588	0.0847	0.0854	1243.84	1.1697
$\sigma$	0.2	0.1413	0.1947	0.2444	0.1942	0.1954	1032.76	1.3754
$\mu$	0.0	-0.1345	0.0052	0.1431	0.0054	0.0051	62257.45	2.5245
$\phi$	0.9	0.8356	0.8958	0.9391	0.8937	0.8989	2253.49	1.3988
$\sigma$	0.2	0.1563	0.2018	0.2575	0.2032	0.1998	1277.12	1.5802
N sampler								
$\mu$	1.0	0.9335	1.0051	1.0742	1.0049	1.0053	9832.80	-
$\phi$	0.9	0.8032	0.8934	0.9493	0.8891	0.8995	880.20	-
$\sigma$	0.1	0.0752	0.1064	0.1468	0.1076	0.1045	489.86	-
$\mu$	1.0	0.9659	1.0012	1.0364	1.0012	1.0012	3422.91	-
$\phi$	0.1	-0.3000	0.0882	0.4591	0.0871	0.0891	1063.36	-
$\sigma$	0.2	0.1413	0.1946	0.2439	0.1942	0.1951	750.85	-
$\mu$	0.0	-0.1344	0.0056	0.1430	0.0055	0.0056	24661.06	-
$\phi$	0.9	0.8354	0.8957	0.9393	0.8937	0.8990	1611.06	-
$\sigma$	0.2	0.1558	0.2019	0.2571	0.2031	0.2002	808.18	-

Table 1: Posterior estimates averaged over 100 simulated data sets assuming a Gaussian  $SCAR(\mu, \phi, \sigma)$  copula model for different parameter settings using the coarse grid (CG) and the naïve (N) sampler, respectively. (Posterior estimates are based on 100 000 MCMC iterations with burnin 5 000 and thinning of 50. The effective sample size (EFS) based on 10 0000 MCMC iterations for both samplers and its ratio  $\text{EFS}_{cg}/\text{EFS}_N$  are given in the last three columns).

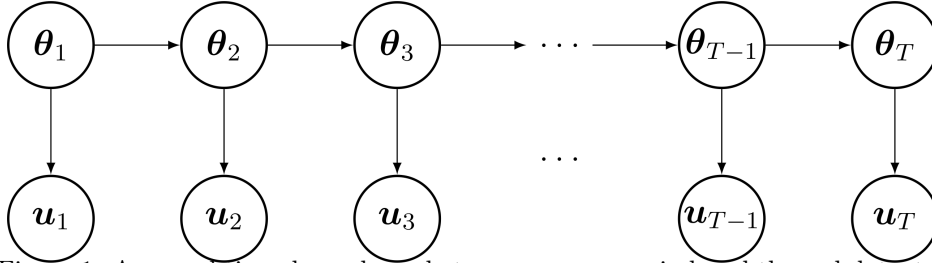


Figure 1: Assumed time dependence between  $u_1, \dots, u_T$  induced through latent variables  $\theta_1, \dots, \theta_T$  (arrows indicate dependence).

CG sampler								
	True	2.5%	50%	97.5%	mean	mode	EFS	$\frac{\text{EFS}_{cg}}{\text{EFS}_N}$
$\mu$	1.0	0.9260	0.9973	1.0657	0.9971	0.9978	31325.61	3.1999
$\phi$	0.9	0.7842	0.8807	0.9392	0.8758	0.8877	1113.03	1.3650
$\sigma$	0.1	0.0748	0.1082	0.1509	0.1095	0.1062	692.68	1.5571
$\mu$	1.0	0.9599	0.9946	1.0300	0.9947	0.9946	4192.09	1.3324
$\phi$	0.1	-0.3423	0.0799	0.4932	0.0803	0.0788	1065.33	1.1605
$\sigma$	0.2	0.1190	0.1800	0.2342	0.1791	0.1811	772.58	1.3443
$\mu$	0.0	-0.1320	0.0065	0.1441	0.0066	0.0066	65627.31	2.3678
$\phi$	0.9	0.8401	0.8965	0.9384	0.8947	0.8992	2617.63	1.4397
$\sigma$	0.2	0.1591	0.2022	0.2536	0.2033	0.2005	1454.21	1.5944
N sampler								
$\mu$	1.0	0.9261	0.9973	1.0664	0.9972	0.9976	9789.47	-
$\phi$	0.9	0.7847	0.8802	0.9395	0.8757	0.8869	815.42	-
$\sigma$	0.1	0.0748	0.1082	0.1509	0.1095	0.1065	444.87	-
$\mu$	1.0	0.9597	0.9949	1.0301	0.9949	0.9949	3146.34	-
$\phi$	0.1	-0.3315	0.0828	0.4977	0.0835	0.0864	918.03	-
$\sigma$	0.2	0.1201	0.1807	0.2342	0.1800	0.1819	574.71	-
$\mu$	0.0	-0.1328	0.0065	0.1435	0.0065	0.0069	27716.70	-
$\phi$	0.9	0.8403	0.8967	0.9385	0.8949	0.8994	1818.20	-
$\sigma$	0.2	0.1592	0.2021	0.2537	0.2033	0.2001	912.08	-

Table 2: Posterior estimates averaged over 100 simulated data sets assuming a double Clayton  $SCAR(\mu, \phi, \sigma)$  copula model for different parameter settings using the coarse grid (CG) and the naïve (N) sampler, respectively. (Posterior estimates are based on 100 000 MCMC iterations with burnin 5 000 and thinning of 50. The effective sample size (EFS) based on 10 000 MCMC iterations for both samplers and its ratio  $\text{EFS}_{cg}/\text{EFS}_N$  are given in the last three columns).

CG sampler								
	True	2.5%	50%	97.5%	mean	mode	EFS	$\frac{\text{EFS}_{cg}}{\text{EFS}_N}$
$\mu$	1.0	0.9679	1.0003	1.0316	1.0001	1.0005	751.86	1.0332
$\phi$	-0.9	-0.9456	-0.8858	-0.7834	-0.8802	-0.8936	876.04	1.3759
$\sigma$	0.1	0.0679	0.1035	0.1503	0.1050	0.1014	546.05	1.5729
$\mu$	1.0	0.9646	1.0030	1.0397	1.0029	1.0030	2822.08	1.1898
$\phi$	-0.1	-0.5101	-0.0239	0.4618	-0.0234	-0.0247	853.38	1.2160
$\sigma$	0.2	0.1016	0.1873	0.2512	0.1848	0.1902	490.05	1.3164
$\mu$	0.0	-0.0436	-0.0004	0.0427	-0.0004	-0.0004	1448.89	1.0214
$\phi$	-0.9	-0.9374	-0.8879	-0.8235	-0.8859	-0.8918	2059.20	1.3390
$\sigma$	0.2	0.1585	0.2158	0.2798	0.2166	0.2122	1188.54	1.4925
N sampler								
$\mu$	1.0	0.9674	0.9997	1.0311	0.9997	0.9999	727.67	-
$\phi$	-0.9	-0.9469	-0.8881	-0.7865	-0.8825	-0.8959	636.71	-
$\sigma$	0.1	0.0667	0.1018	0.1486	0.1033	0.0992	347.16	-
$\mu$	1.0	0.9634	1.0000	1.0397	1.0005	0.9996	2371.97	-
$\phi$	-0.1	-0.5207	-0.0246	0.4618	-0.0255	-0.0215	701.78	-
$\sigma$	0.2	0.0997	0.1794	0.2506	0.1784	0.1811	372.27	-
$\mu$	0.0	-0.0434	-0.0005	0.0420	-0.0005	-0.0005	1418.54	-
$\phi$	-0.9	-0.9386	-0.8957	-0.8373	-0.8937	-0.8991	1537.90	-
$\sigma$	0.2	0.1565	0.2028	0.2585	0.2041	0.2009	796.32	-

Table 3: Posterior estimates averaged over 100 simulated data sets assuming a double Gumbel  $SCAR(\mu, \phi, \sigma)$  copula model for different parameter settings using the coarse grid (CG) and the naïve (N) sampler, respectively. (Posterior estimates are based on 100 000 MCMC iterations with burnin 5 000 and thinning of 50. The effective sample size (EFS) based on 10 000 MCMC iterations for both samplers and its ratio  $\text{EFS}_{cg}/\text{EFS}_N$  are given in the last three columns).

	Dow Jones	NASDAQ	S&P500	DAX
Initial Date	1990/03/23	1990/03/23	2002/06/06	2002/06/06
Final Date	2000/03/26	2000/03/26	2009/11/06	2009/11/06
$\omega$	5.456e-07 (1.422e-07)	2.460e-06 (4.038e-07)	1.051e-06 (2.850e-07)	1.875e-06 (5.562e-07)
$\alpha$	4.674e-02 (4.029e-03)	1.005e-01 (7.748e-03)	7.026e-02 (1.016e-02)	8.356e-02 (1.190e-02)
$\beta$	9.482e-01 (4.995e-03)	8.829e-01 (9.418e-03)	9.212e-01 (1.071e-02)	9.094e-01 (1.179e-02)

Table 4: Parameter estimates of marginal GARCH(1,1) models with Gaussian innovations together with estimated standard deviation in parentheses using maximum likelihood for each margin.

	Dow Jones vs. NASDAQ					S&P500 vs. DAX				
	2.5%	50%	97.5%	mean	mode	2.5%	50%	97.5%	mean	mode
Gaussian-SCAR										
$\mu$	0.5361	0.5849	0.6358	0.5854	0.5848	0.3830	0.4238	0.4615	0.4239	0.4227
$\phi$	0.9148	0.9584	0.9851	0.9556	0.9620	0.4592	0.6849	0.8435	0.6724	0.6969
$\sigma$	0.0277	0.0469	0.0709	0.0479	0.0459	0.1054	0.1675	0.2289	0.1666	0.1708
Clayton-SCAR										
$\mu$	0.4490	0.4863	0.5193	0.4854	0.4870	0.3220	0.3544	0.3873	0.3547	0.3541
$\phi$	0.5395	0.7646	0.8730	0.7432	0.7894	0.1580	0.4921	0.7323	0.4745	0.5211
$\sigma$	0.1005	0.1459	0.2165	0.1505	0.1406	0.1094	0.1663	0.2229	0.1666	0.1672
Gumbel-SCAR										
$\mu$	0.4953	0.5362	0.5742	0.5359	0.5355	0.3301	0.3725	0.4127	0.3725	0.3719
$\phi$	0.7473	0.9036	0.9584	0.8861	0.9192	0.6269	0.8001	0.9016	0.7885	0.8167
$\sigma$	0.0477	0.0805	0.1393	0.0849	0.0747	0.0797	0.1264	0.1806	0.1272	0.1251

Table 5: Estimated posterior quantiles, mean and mode for the time varying copula parameters using a Gaussian, Clayton and Gumbel SCAR copula model based on 1 900 recorded MCMC iterations for the two data sets

	Dow Jones vs. NASDAQ			S&P500 vs. DAX		
	RW50	RW125	RW250	RW50	RW125	RW250
Gaussian-SCAR	<b>14.22</b>	<b>12.40</b>	12.51	17.45	16.01	14.44
Clayton-SCAR	25.94	20.56	16.58	16.20	<b>10.95</b>	<b>8.05</b>
Gumbel-SCAR	15.36	12.72	<b>12.28</b>	<b>14.04</b>	12.83	10.93

Table 6: Sum of square of distances between SCAR based Kendall's  $\tau_t$  estimates and empirical ones based on rolling windows of size 50, 125 and 250 for the two data sets.

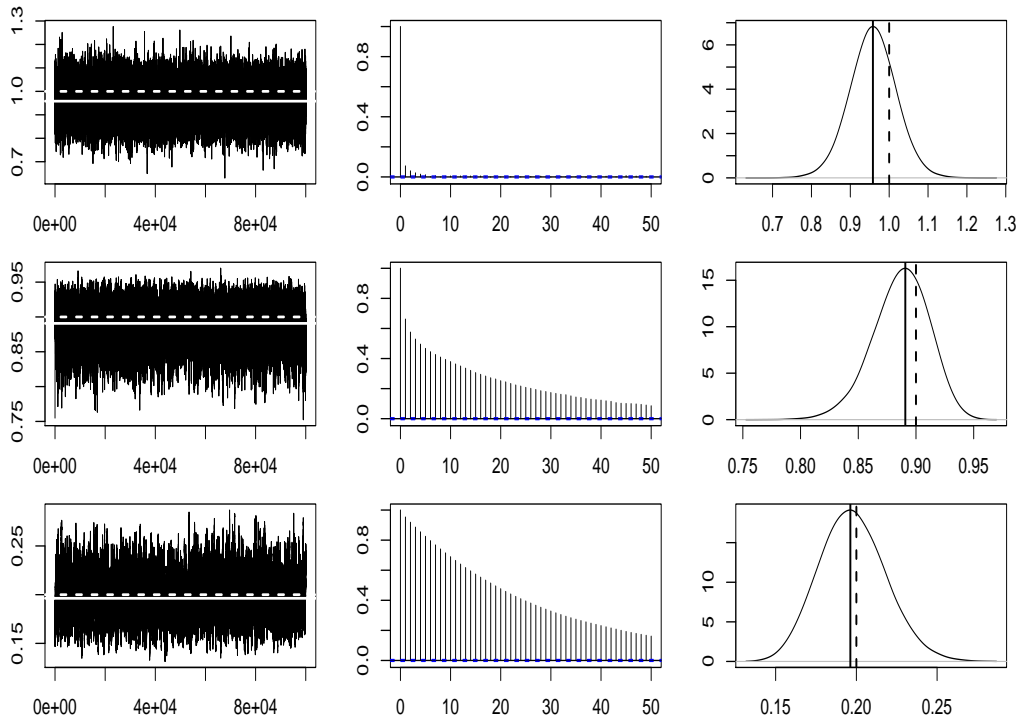


Figure 2: MCMC traces, estimated autocorrelations and posterior densities of  $\alpha = (\mu, \phi, \sigma)^\top$  for a bivariate data set with  $T = 1000$  from a Gaussian  $SCAR(1, .9, .2)$  copula model using the CG sampler (rows correspond to  $\mu, \phi$  and  $\sigma$ , respectively and — MAP estimates, - - - true values)

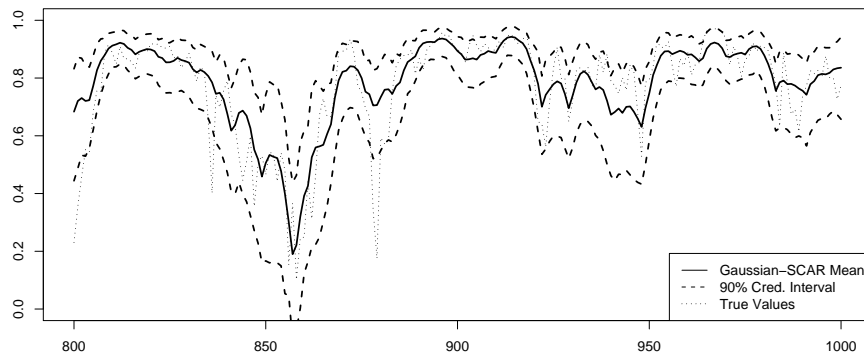


Figure 3: Estimated posterior Kendall's  $\tau_t$  for the last 200 observations from a Gaussian  $SCAR(1, .9, .2)$  copula model of size  $T = 1000$  using the CG sampler



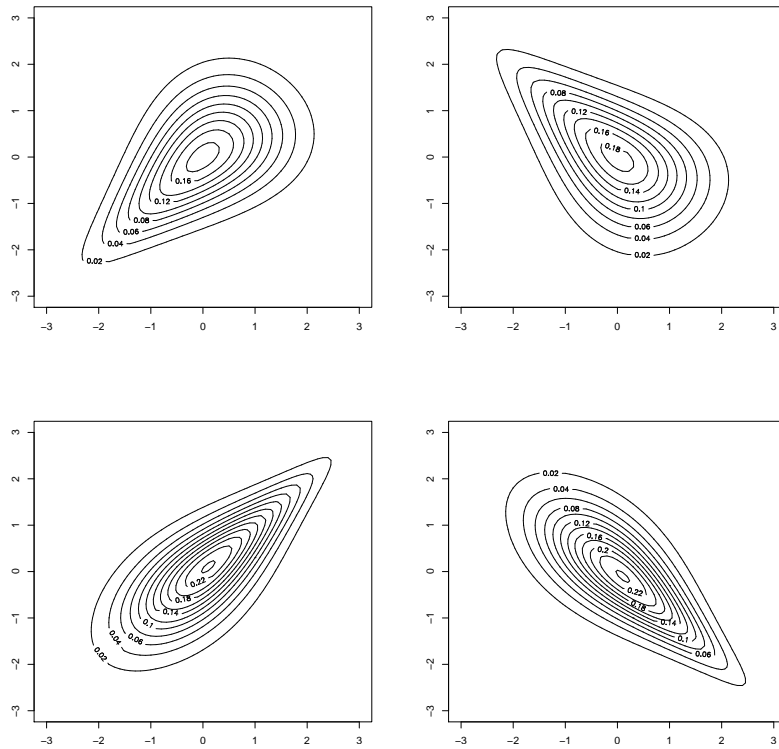


Figure 4: First row: contours for a double Clayton ( $DC(\theta)$ ) copula with standard normal margins with  $\theta = 1$  (left panel) and for  $\theta = -1$  (right panel), respectively. Second row: contours for a double Gumbel ( $DG(\theta)$ ) copula with standard normal margins with  $\theta = 2$  (left panel) and for  $\theta = -2$  (right panel), respectively.

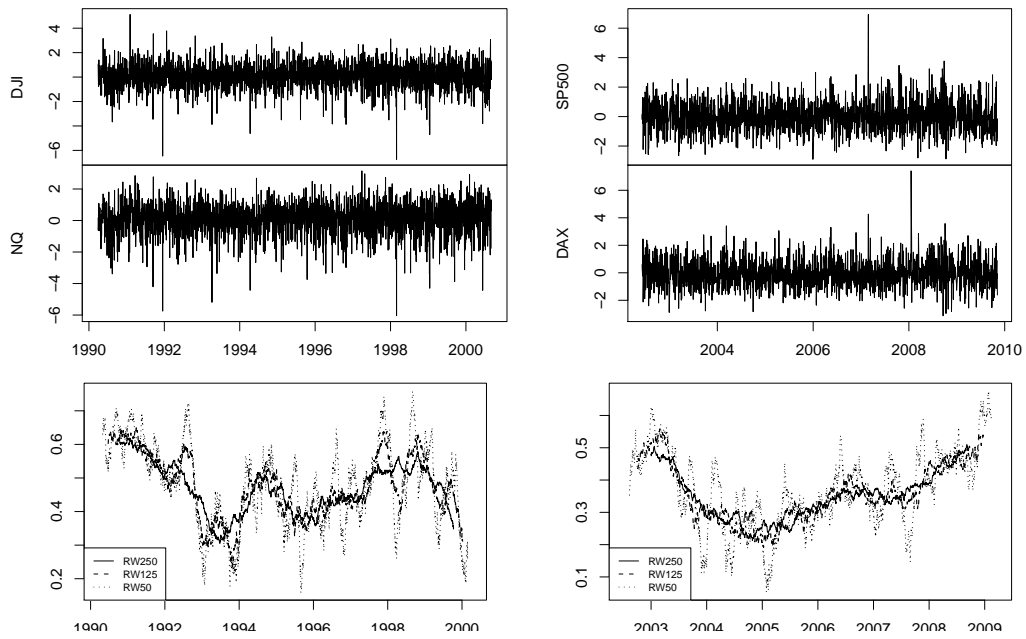


Figure 5: Top row: Standardised innovations for the log returns for the DJI and the NASDAQ data from March 23th, 1990 to March 26th, 2000 after GARCH(1,1) filtering (left panel) and for the S&P500 and DAX data from June 6th, 2002 to November 11th, 2009 (right panel). Bottom row: Rolling windows estimates for time varying Kendall's  $\tau_t$  by using a window size of 250, 125 and 50 values, respectively, for both data sets.

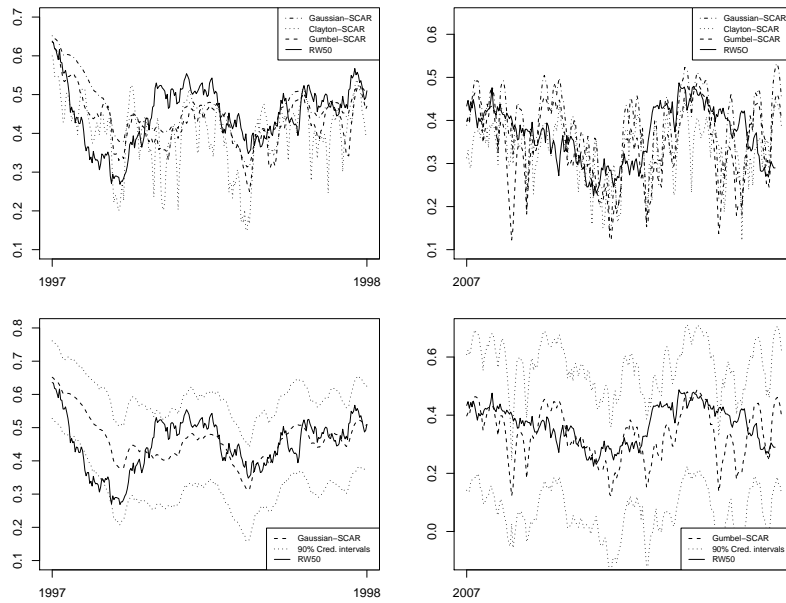


Figure 6: Top row: Rolling window based with window size 50 and estimated posterior Kendall's  $\tau_t$  based on Gaussian, Clayton and Gumbel SCAR models for the DJI-NASDAQ data in 1997 (left panel) and for the S&P500-DAX data for 2007 (right panel) Bottom row: Point-wise 90% credible intervals for Kendall's  $\tau_t$  for the chosen model (Gaussian SCAR (Gumbel SCAR) for DJI-NASDAQ (S&P500-DAX) data, respectively).

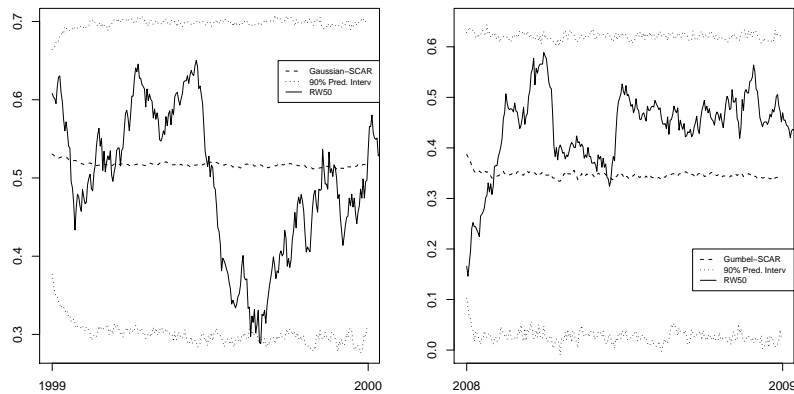


Figure 7: Pointwise 90% predictive intervals for Kendall's  $\tau_t$  in 1999 based on Gaussian SCAR model for the DJI-NASDAQ data (left panel) and in 2008 based on Gumbel SCAR for the S&P500-DAX data (right panel).