Bayesian Inference for Pair-copula Constructions of Multiple Dependence

Claudia Czado and Aleksey Min
Technische Universität München
cczado@ma.tum.de, aleksmin@ma.tum.de

January 29, 2008

Overview

1. Introduction
2. Pair-copula constructions (PCC)
3. Bayesian Analysis of PCC's
4. Application: Financial Returns
5. Bayesian Model Selection among PCC's
6. Euro Swap Rates
7. Summary and Outlook

Introduction

- Bedford and Cooke (2001) and Bedford and Cooke (2002) gave a probabilistic construction of multivariate distributions based on simple building blocks called pair-copulas. See also Kurowicka and Cooke (2006).
- Extends work by Joe (1996).
- Aas, Czado, Frigessi, and Bakken (2007) used the PCC construction to construct flexible multivariate copulas based on pair-copulas such as bivariate Gaussian, t-, Gumbel and Clayton copulas

Introduction

- Parameters are estimated by maximum likelihood (ML) and PCC's are sucessfully applied to financial return data
- Berg and Aas (2007) give further example of PCC's and developed an **R**-package for fitting PCC's using MLE
- The Fisher information is difficult to handle analytically, so therefore estimated standard errors for parameter estimates are so far absent
- Here we follow a Bayesian approach where such interval estimates are easy to establish

Multivariate Distributions

Consider n random variables $\mathbf{X} = (X_1, \ldots, X_n)$ with

- joint density $f(x_1, \ldots, x_n)$ and marginal densities $f_i(x_i), i = 1, \ldots, n$
- joint cdf $F(x_1, \ldots, x_n)$ and marginal cdf's $F_i(x_i), i = 1, \ldots, n$
- factorization

$$f(x_1, \ldots, x_n) = f_n(x_n) \cdot f(x_{n-1}|x_n) \cdot f(x_{n-2}|x_{n-1}, x_n) \ldots \cdot f(x_1|x_2, \ldots, x_n) \quad (1)$$

Copula

A copula is a multivariate distribution on $[0,1]^n$ with uniformly distributed marginals.

- copula cdf $C(u_1, \ldots, u_n)$
- copula density $c(u_1, \ldots, u_n)$

Using Sklar's Theorem (1959) we have for absolutely continuous distributions with continuous marginal cdf's

$$f(x_1, \ldots, x_n) = c_{12\cdots n}(F_1(x_1), \ldots F_n(x_n)) \cdot f_1(x_1) \cdots f_n(x_n) \qquad (2)$$

for some n-variate copula density $c_{12\cdots n}(\cdot)$.

Pair-copula constructions (PCC)

- $n = 3$

$$\begin{aligned} f_{1|23}(x_1|x_2, x_3) &= \frac{f_{12|3}(x_1, x_2|x_3)}{f_{2|3}(x_2|x_3)} \\ &= c_{12|3}(F_{1|3}(x_1|x_3), F_{2|3}(x_2|x_3)) \cdot f_{1|3}(x_1|x_3) \end{aligned}$$

- general n For $\mathbf{v} = (v_1, \ldots, v_d)$ and any $j = 1, \ldots, d$

$$\begin{aligned} f(x|\mathbf{v}) &= c_{xv_j|\mathbf{v}_{-j}}(F(x|\mathbf{v}_{-j}), F(v_j|\mathbf{v}_{-j})) \cdot f(x|\mathbf{v}_{-j}) \\ \mathbf{v}_{-j} &= (v_1, \ldots, v_{j-1}, v_{j+1}, \ldots, v_d) \\ c_{xv_j|\mathbf{v}_{-j}}(\cdot) &= \text{bivariate copula density} \end{aligned}$$

- Combining this decomposition of the conditional distribution with the factorization (1), we derive at a decomposition of $f(x_1, \ldots, x_n)$ that only consist of pair-copulae.
- We call this a pair-copula construction (PCC).

Conditional cdf's

- Univariate v: Since $f(x|v) = c_{xv}(F_x(x), F_v(v))f_x(x)$ we have

$$\begin{aligned} F(x|v) &= \int_{-\infty}^{x} c_{xv}(F_x(u), F_v(v))f_x(u)du \\ &= \frac{\partial\, C_{xv}(F_x(x), F_v(v))}{\partial F_v(v)} \end{aligned}$$

- General v: Under regularity conditions Joe (1996) showed that

$$F(x|\mathbf{v}) = \frac{\partial\, C_{x,v_j|\mathbf{v}_{-j}}(F(x|\mathbf{v}_{-j}), F(v_j|\mathbf{v}_{-j}))}{\partial F(v_j|\mathbf{v}_{-j})},$$

Vines

- For high-dimensional distributions there are many possible pair-copula constructions.
- Bedford and Cooke (2001) introduced a graphical model called regular vine to help organize them.
- The class of regular vines is large and embraces a large number of possible PCC's.
- We concentrate on two special cases (Kurowicka and Cooke 2004):
  - D-vine
  - Canonical Vine

Vines(2)

- An n-dimensional vine is represented by n-1 trees.
- Tree j has $n+1-j$ nodes and $n-j$ edges.
- Each edge corresponds to a pair-copula density.
- Edges in tree j become nodes in tree $j+1$.
- Two nodes in tree $j+1$ are joined by an edge if the corresponding edges in tree j share a node.
- The complete decomposition is defined by the $\frac{n(n-1)}{2}$ edges (i.e. pair copula densities) and the marginal densities.

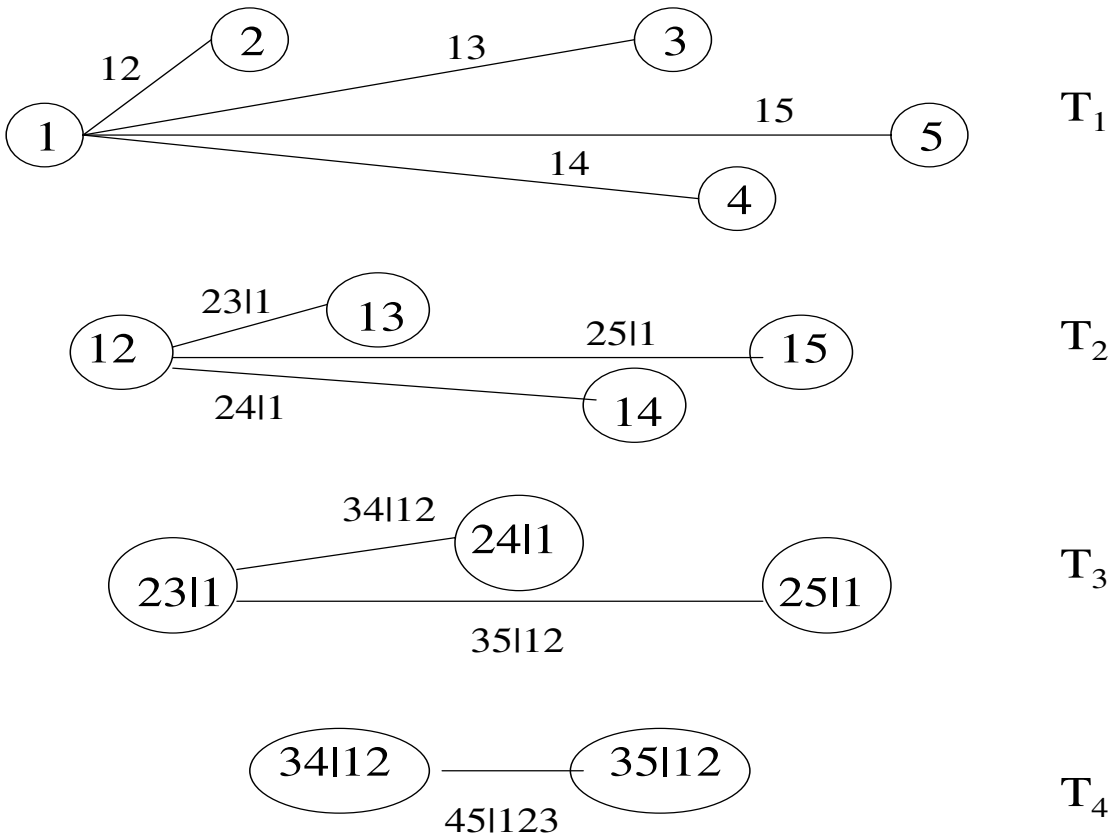
Canonical and D-vines

- Canonical vine:
  A regular vine for which each tree has a unique node that is connected to $n - j$ edges.
- D-vine:
  A regular vine for which no node in any tree is connected to more than two edges
- Abbreviations:

$$\begin{aligned} c_{ij|\mathbf{v}} &:= c_{ij|\mathbf{v}}(F_{i|\mathbf{v}}(x_i|\mathbf{x_v}), F_{j|\mathbf{v}}(x_j|\mathbf{x_v})) \\ f_j &:= f_j(x_j) \\ f_{\mathbf{v}} &:= f_{\mathbf{v}}(x_{\mathbf{v}}) \end{aligned}$$

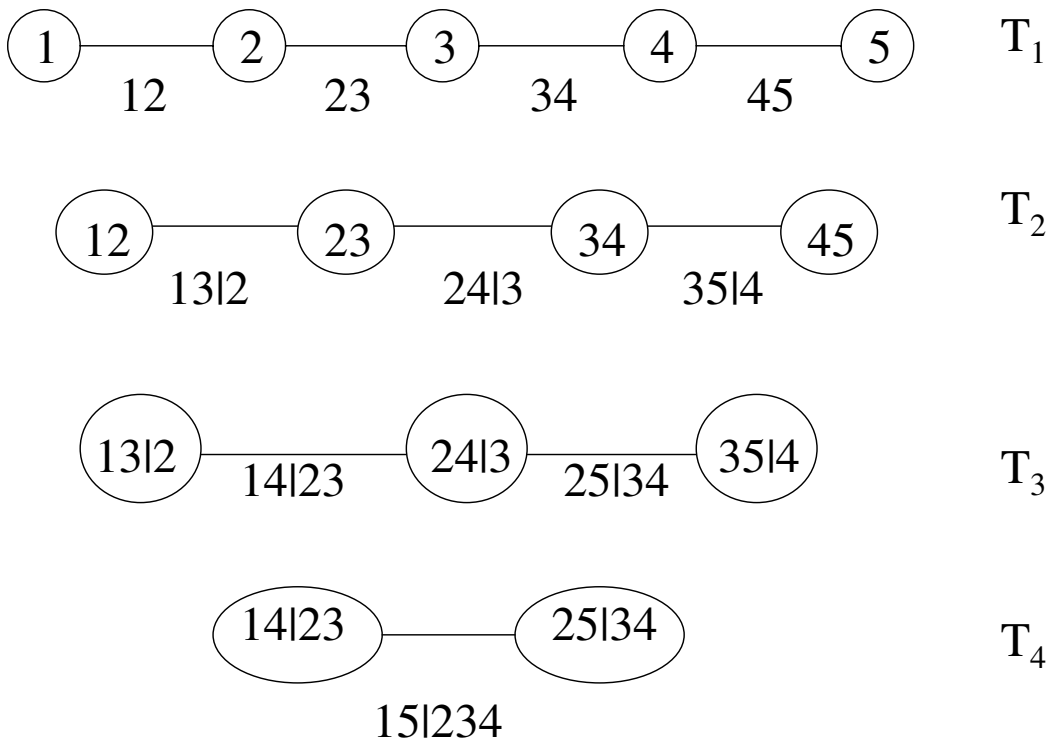
Five dimensional canonical vine

$$f_{12345} = f_1 \cdot f_2 \cdot f_3 \cdot f_4 \cdot f_5 \cdot c_{12} \cdot c_{13} \cdot c_{14} \cdot c_{15} \cdot c_{23|1} \cdot c_{24|1} \cdot c_{25|1} \cdot c_{34|12} \cdot c_{35|12} \cdot c_{45|123}$$



Five dimensional D-vine

$$f_{12345} = f_1 \cdot f_2 \cdot f_3 \cdot f_4 \cdot f_5 \cdot c_{12} \cdot c_{23} \cdot c_{34} \cdot c_{45} \cdot c_{13|2} \cdot c_{24|3} \cdot c_{35|4} \cdot c_{14|23} \cdot c_{25|34} \cdot c_{15|234}$$



General density expressions

- Canonical vine density

$$\prod_{k=1}^{n} f_k \prod_{j=1}^{n-1} \prod_{i=1}^{n-j} c_{j,j+i|1,\ldots,j-1}$$

- D-vine density

$$\prod_{k=1}^{n} f_k \prod_{j=1}^{n-1} \prod_{i=1}^{n-j} c_{i,i+j|i+1,\ldots,i+j-1}$$

where index j identifies the trees, while i runs over the edges in each tree.

Bayesian Analysis of PCC's with t-copula pairs

Assume bivariate t-copula for each pair copula term, i.e.

$$\boldsymbol{\theta} = (\rho, \nu) \ \rho = \textbf{correlation} \ , \nu = \ \textbf{df}$$

$$c(u_1, u_2|\boldsymbol{\theta}) = \frac{\Gamma(\frac{\nu+2}{2})/\Gamma(\frac{\nu}{2})}{\nu\pi t_\nu(x_1)t_\nu(x_2)\sqrt{1-\rho^2}}\left(1+\frac{x_1^2+x_2^2-2\rho\, x_1\, x_2}{\nu(1-\rho^2)}\right)^{-\frac{\nu+2}{2}},$$

where

$$x_1 = t_\nu^{-1}(u_1)$$
$$x_2 = t_\nu^{-1}(u_2)$$
$$t_\nu(\cdot) = \text{pdf of univariate } t_\nu \text{ distribution}$$
$$t_\nu^{-1}(\cdot) = \text{quantile function of } t_\nu \text{ distribution}$$

Bayesian Analysis of PCC's with t-copula pairs

- For the PCC's we assume a regular vine structure with edge indices denoted by $s := ij|\mathbf{v}$ for $s \in \mathcal{S}$.
- Let $\boldsymbol{\theta}_s = (\rho_s, \nu_s)$ the corresponding parameters of the pair copula associated with edge s.
- Let $\boldsymbol{\theta} = \{\boldsymbol{\theta}_s, s \in \mathcal{S}\}$ denote the vector of all parameters of the PCC with t-copula pairs
- We assume uniform priors for ρ_s on $(-1, 1)$ and for ν_s on $(1, U)$ for each s and prior independence among all parameters
- Inference is based on the posterior distribution given by

$$p(\boldsymbol{\theta}|data) = \frac{f(data|\boldsymbol{\theta}) \times p(\boldsymbol{\theta})}{\int f(data|\boldsymbol{\theta}) \times p(\boldsymbol{\theta}) d\boldsymbol{\theta}}$$

where $f(data|\boldsymbol{\theta})$ is the likelihood of the PCC and $p(\boldsymbol{\theta})$ the prior for θ

Bayesian Analysis of PCC's with t-copula pairs

- Since $\boldsymbol{\theta}$ is high dimensional the posterior is not analytically tractable.
- Posterior distribution is estimated using Markov Chain Monte Carlo (MCMC) methods
- For each univariate parameter Metropolis-Hastings (MH) updates with symmetric normal random walk proposals are used
- Proposal variances are determined by pilot runs to achieve acceptance rates between 20% and 80%.
- Posterior mean, mode, median and density are estimated using the MCMC iterates.
- Credible intervals are estimated by empirical quantiles of the MCMC iterates.

Application: Financial Returns

- Tail dependence properties are important in finance.
- *n*-dimensional Student's t-copula has been often used for modeling financial returns, however it only has a single parameter for tail dependence.
- Pair-copula constructions allows for multiple parameters for modeling tail dependence.

Data set

- Daily data from Jan. 4, 1999 until July 8, 2003 for
  - **T** = Norwegian stock index (TOTX)
  - **M** = MSCI world stock index
  - **B** = Norwegian bond index (BRIX)
  - **S** = SSBWG hedged bond index
- Standardized residuals of a AR(1)-GARCH(1,1) model for the log return $x_{i,t}$

$$\begin{aligned} x_{i,t} &= c_i + \alpha_i\, x_{i,t-1} + \sigma_{i,t}\, z_{i,t}, \\ \mathrm{E}[z_{t,i}] &= 0 \text{ and } \mathrm{Var}[z_{t,i}] = 1, \\ \sigma^2_{i,t} &= a_{i,0} + a_i\, \epsilon^2_{i,t-1} + b_i\, \sigma^2_{i,t-1} \text{ where } \epsilon_{i,t-1} = \sigma_{i,t}\, z_{i,t} \end{aligned}$$

are found to be independent and are transformed to uniform margins

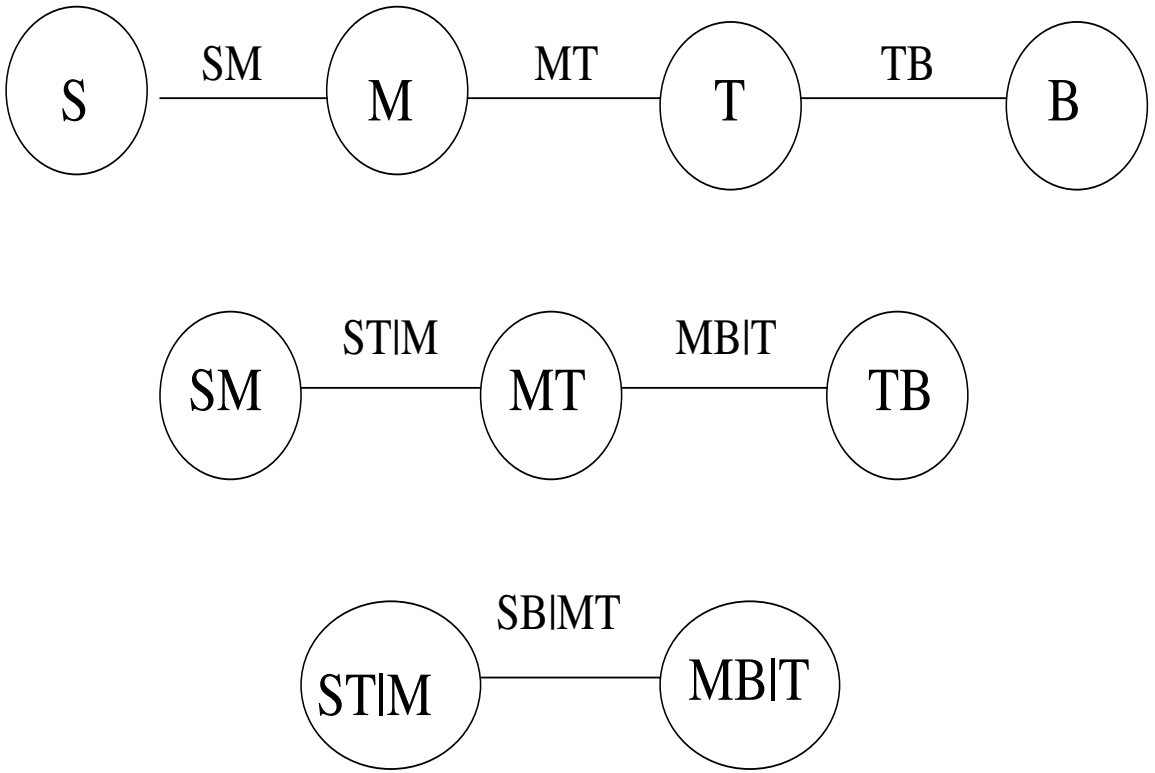
Data set

- Fitted degree of freedom for a t-copula to each pair

Between	M	T	B
S	4.21	34.16	14.47
M		16.65	15.48
T			12.60

Which PCC?

Strongest dependence between S and M, M and T and T and B, which will be used in the top tree of a D-vine



Bayesian Inference of the Norwegian return data

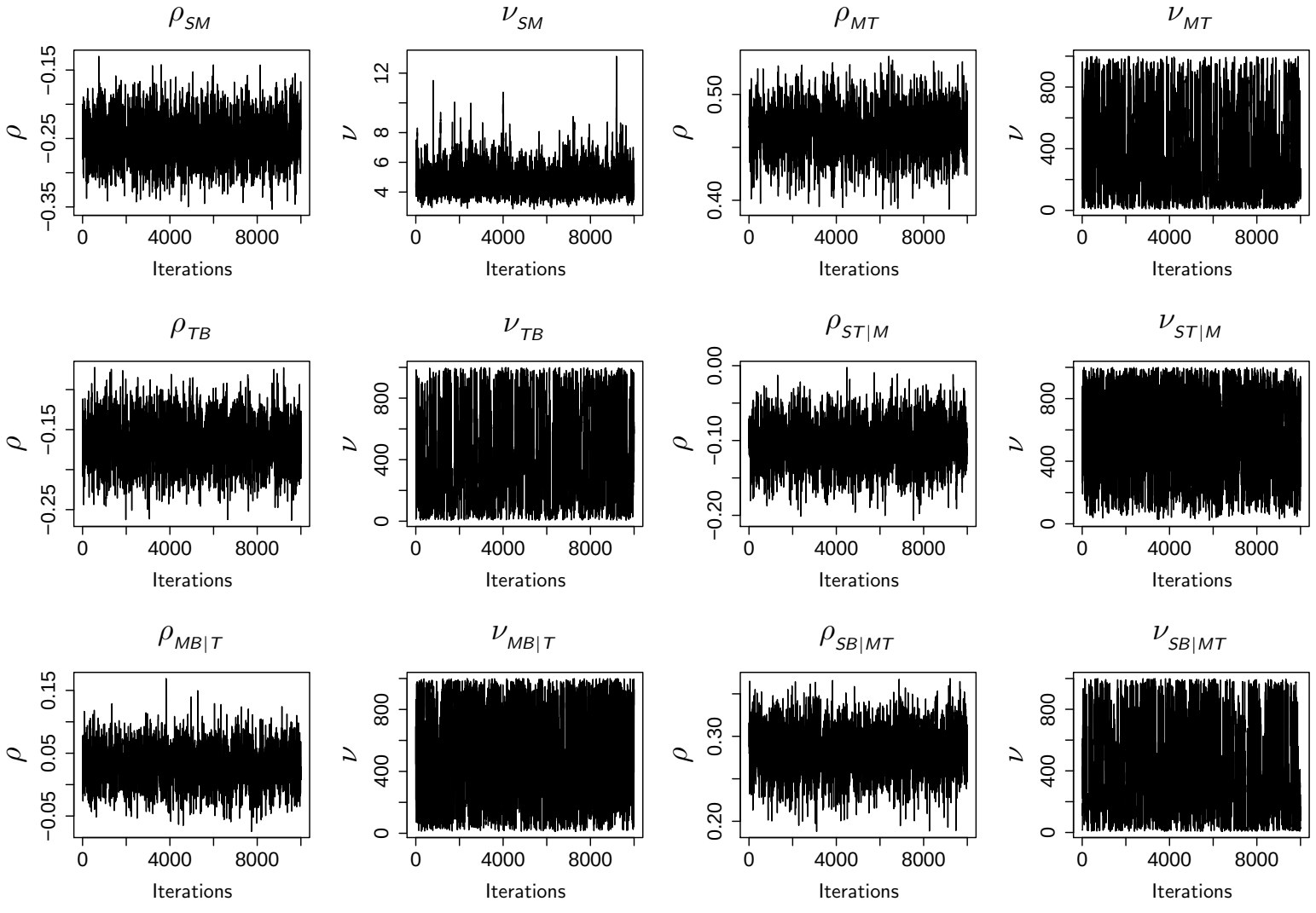
- For each pair copula we assume a bivariate t-copula with correlation ρ and ν degree of freedom.
- For each ρ we assume uniform(-1,1) prior and for ν a uniform(1,1000) prior. Further prior independence between all parameters are assumed.
- For the likelihood we assume the following PCC

$$\begin{aligned} c(x_S, x_M, x_T, x_B) &= c_{SM}\left(F(x_S), F(x_M)\right) \cdot c_{MT}\left(F(x_M), F(x_T)\right) \\ &\cdot\ c_{TB}\left(F(x_T), F(x_B)\right) \cdot c_{ST|M}\left(F(x_S|x_M), F(x_T|x_M)\right) \\ &\cdot\ c_{MB|T}\left(F(x_M|x_T), F(x_B|x_T)\right) \\ &\cdot\ c_{SB|MT}\left(F(x_S|x_M, x_T), F(x_B|x_M, x_T)\right) \end{aligned}$$

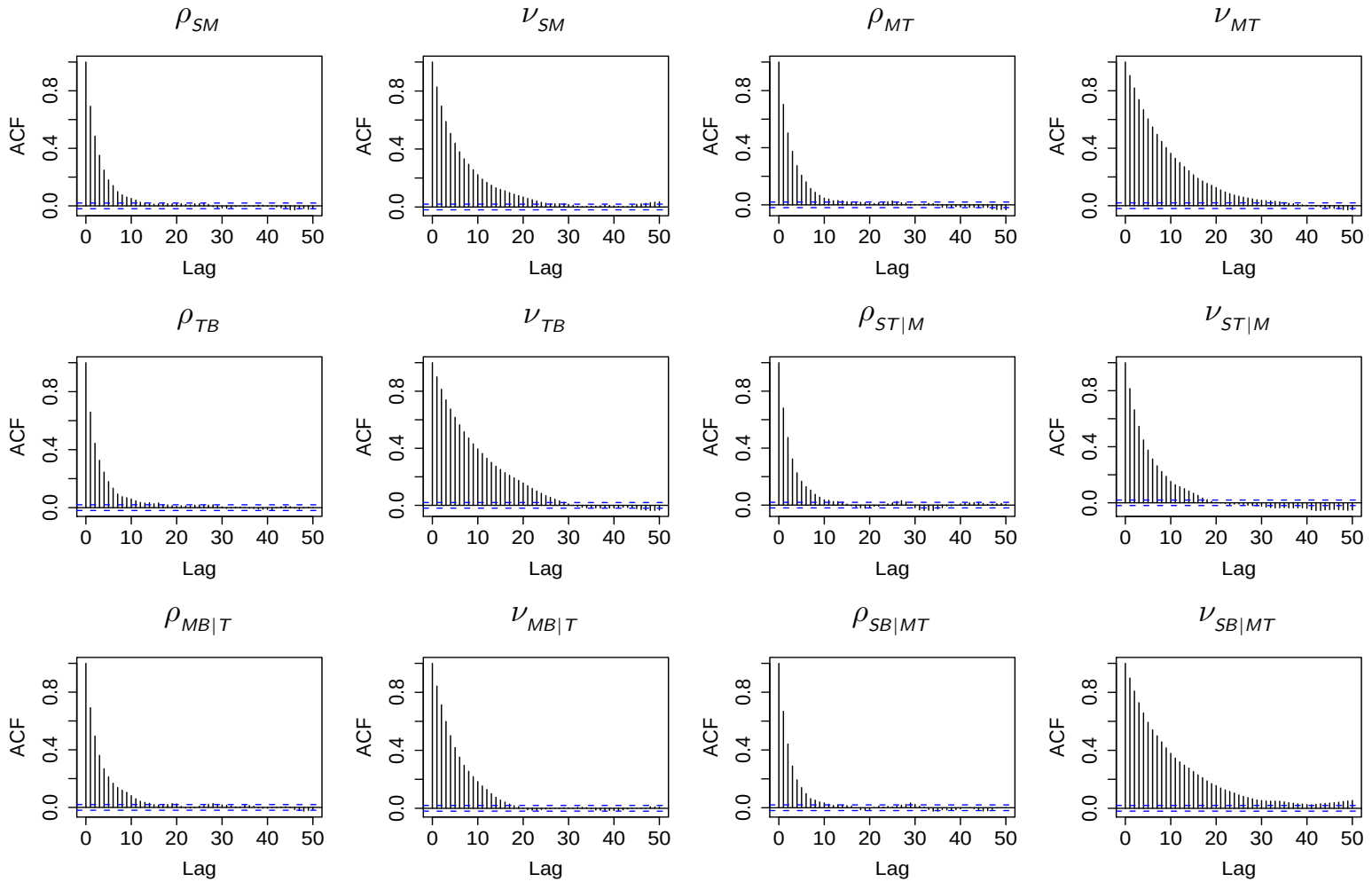
Bayesian Inference of the Norwegian return data

- For the MH updates in the MCMC algorithm we used
  - symmetric random walk proposals
  - normal proposals with variances determined by pilot runs to achieve acceptance rates between 25% and 80%
- 10000 MCMC iterations were run
- MLE estimates were used as starting values
- Implemented using Daniel Berg's **R**-package

Trace Plots of MCMC iterations

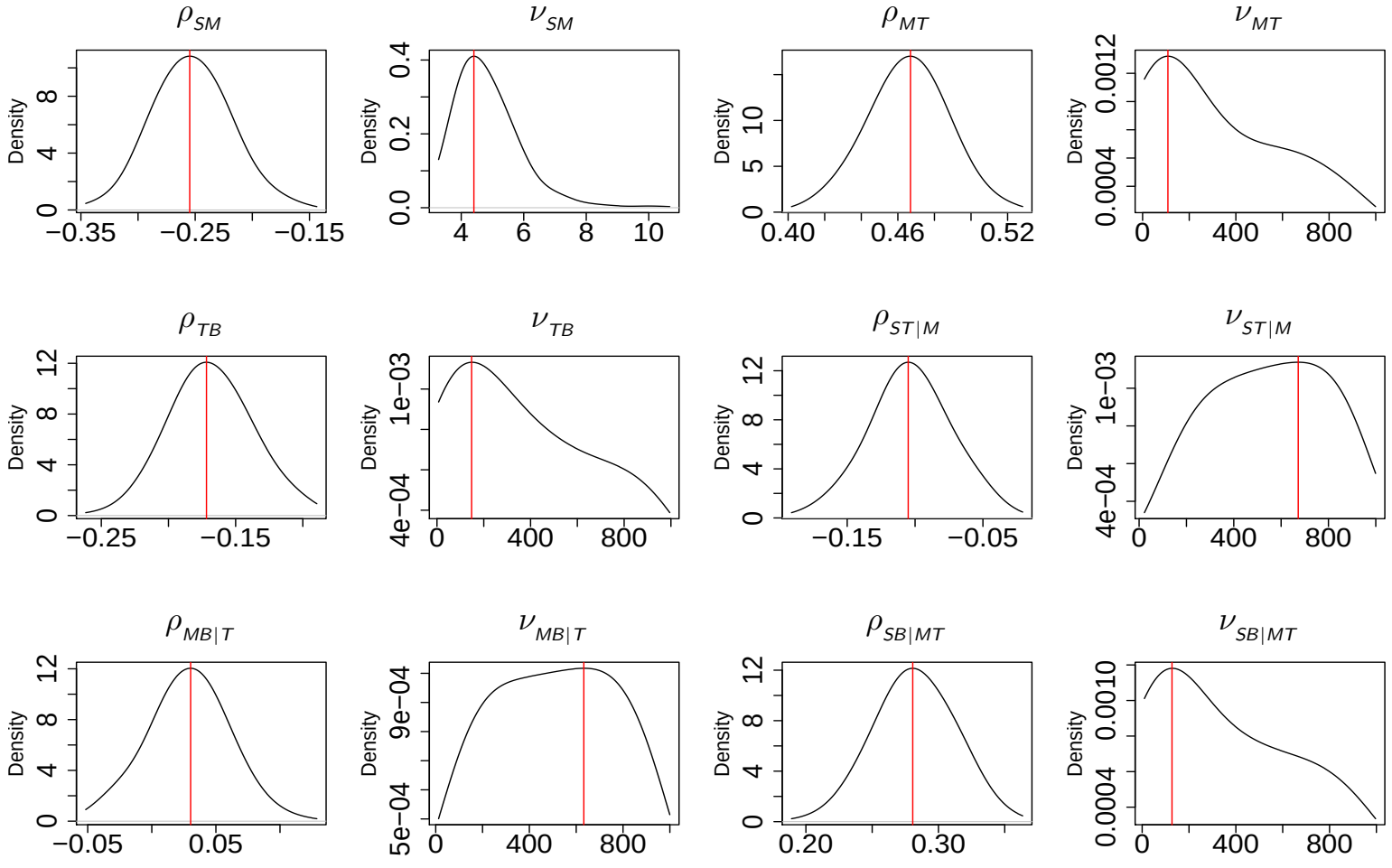


Autocorrelations of the MCMC iterations



For the further analysis we used burnin 1000 iteration and only every 20th iteration.

Estimated Posterior Densities



Summary statistics for thinned MCMC

	2.5% Quantile	50% Quantile	97.5% Quantile	Est. Post. Mean	Est. Post. Mode	MLE
ρ_{SM}	-0.316	-0.254	-0.184	-0.253	-0.25	-0.25
ν_{SM}	3.483	4.658	7.263	4.849	4.40	4.34
ρ_{MT}	0.422	0.466	0.508	0.465	0.47	0.47
ν_{MT}	13.304	228.006	948.340	326.388	108.61	16.26
ρ_{TB}	-0.224	-0.170	-0.106	-0.168	-0.17	-0.17
ν_{TB}	12.829	321.556	973.567	383.984	148.08	13.17
$\rho_{ST\|M}$	-0.163	-0.104	-0.047	-0.103	-0.11	-0.11
$\nu_{ST\|M}$	106.098	560.460	964.869	550.005	672.33	300.00
$\rho_{MB\|T}$	-0.033	0.031	0.090	0.029	0.03	0.03
$\nu_{MB\|T}$	45.736	514.904	972.857	510.722	632.56	45.59
$\rho_{SB\|MT}$	0.226	0.281	0.337	0.282	0.28	0.28
$\nu_{SB\|MT}$	13.557	285.269	958.333	366.153	127.36	15.04

red corresponds to uncorrelatedness

magenta corresponds to near normality

Model fit: Kendall's τ

- Kendall's τ measures dependence and is defined by

$$\tau := P((X_1 - X_2)(Y_1 - Y_2) > 0) - P((X_1 - X_2)(Y_1 - Y_2) < 0)$$

  where (X_1, Y_1) and (X_2, Y_2) are i.i.d.
- τ is invariant under monotone transformations
- Relationship between τ and ρ for bivariate t-distributions

$$\tau = \frac{2}{\pi} \arcsin(\rho)$$

- Estimated posterior summaries of τ

	2.5% Quantile	50% Quantile	97.5% Quantile	Est. Post. Mean	Est. Post Mode	Empirical τ
τ_{SM}	-0.204	-0.164	-0.1179	-0.163	-0.164	-0.158
τ_{MT}	0.277	0.308	0.3393	0.308	0.309	0.313
τ_{TB}	-0.144	-0.109	-0.0677	-0.108	-0.110	-0.110

Model fit: Tail dependence

- Lower and upper tail dependence:

$$\lambda_l := \lim_{v \to 0} P(X \leq F_X^{-1}(v) | Y \leq F_Y^{-1}(v))$$
$$\lambda_u := \lim_{v \to 1} P(X \geq F_X^{-1}(v) | Y \geq F_Y^{-1}(v))$$

- For t distributions we have

$$\lambda := \lambda_l = \lambda_u = 2\, t_{\nu+1}\left(-\sqrt{\nu+1}\sqrt{\frac{1-\rho}{1+\rho}}\right)$$

- Estimated posterior summaries of λ

	2.5% Quantile	50% Quantile	97.5% Quantile	Est. Post. Mean	Est. Post Mode	Empirical λ_l
λ_{SM}	0.0048	0.02384	0.04656	0.02426	0.02312	0.02941
λ_{MT}	0.0000	0.00000	0.04378	0.00397	0.00000	0.12500
λ_{TB}	0.0000	0.00000	0.00051	0.00009	0.00000	0.02778

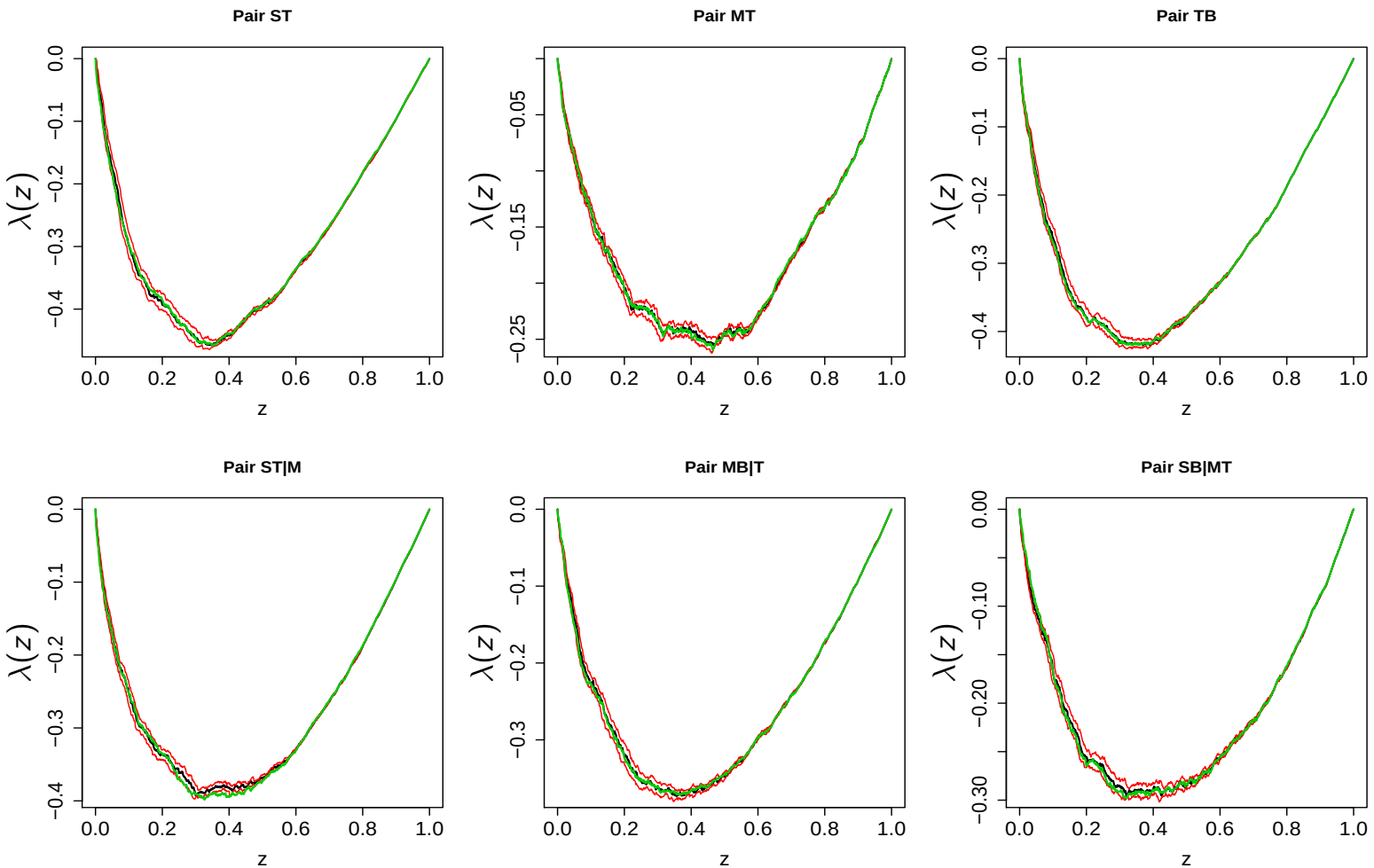
Model fit: Lambda function

- The λ function is defined as a shifted copula distribution function given by

$$\begin{aligned} \lambda(z, \boldsymbol{\theta}) &:= z - K(z, \boldsymbol{\theta}) \\ K(z, \boldsymbol{\theta}) &:= P(C(u_1, u_2, \boldsymbol{\theta}) \leq z) \end{aligned}$$

- Here $\boldsymbol{\theta}$ denotes the parameters of the bivariate copula pair.
- For the conditional pairs $ST|M, MB|T$ and $SB|MT$ an empirical estimate of $\lambda(z, \boldsymbol{\theta})$ is based on transformed data using posterior mode estimates

Estimated pointwise posterior mode $\lambda(z, \theta)$



red = 95% CI, black = est. posterior mode, green = empirical

Bayesian Model Selection among PCC's

Want to select among

- K models $M_1, \cdots, M_K$,
- where Model M_k has parameter $\boldsymbol{\theta}_k$

and compare them on the basis of posterior model probabilities given by

$$P(\text{ Model } M_k|\text{data}), k = 1, \cdots, K$$

We distinguish two situations

- K is small and it is feasible timewise to fit all models (Congdon (2006) and Scott (2002))
- K is large and we want only to fit models which are probable using reversible jump MCMC (RJMCMC) (Green (1995))

Bayesian Model Selection : K small

Congdon (2006) made the following assumptions

- The distribution of the data is independent of $\{\boldsymbol{\theta}_{j\neq k}\}$ given M_k
- Independence among $\boldsymbol{\theta}_k$'s given Model M

and showed that

- Posterior distributions of $\boldsymbol{\theta}_k$ are independent given $M = M_k$ and can be sampled individually
- and used that

$$P(M = M_k|\text{data}, \boldsymbol{\theta}) \propto P(\text{data}|\boldsymbol{\theta}_k, M = M_k)P(\boldsymbol{\theta}|M = M_k)P(M = M_k) \quad (3)$$

Bayesian Model Selection : K small

Assume that K independently MCMC runs result in

$$\begin{array}{lcl} M_1 : \boldsymbol{\theta}_1^{(r)}, r = 1, \cdots, R & & p(\boldsymbol{\theta}_1|\text{data}) \\ \vdots & \text{approximating} & \vdots \\ M_K : \boldsymbol{\theta}_K^{(r)}, r = 1, \cdots, R & & p(\boldsymbol{\theta}_K|\text{data}) \end{array}$$

Using $\{\boldsymbol{\theta}^{(r)} := (\boldsymbol{\theta}_1^{(r)}, \cdots, \boldsymbol{\theta}_K^{(r)}), r = 1, \cdots, R\}$

$$P(M|\text{data}) = \int P(M|\boldsymbol{\theta}, \text{data}) p(\boldsymbol{\theta}|\text{data}) d\boldsymbol{\theta}$$

is approximated by $\widehat{P(M|\text{data})} := \frac{1}{R} \sum_{r=1}^{R} P(M|\boldsymbol{\theta}^{(r)}, \text{data})$.

Bayesian Model Selection : K small

Using (3) we can estimate $P(M = M_k|\text{data}, \boldsymbol{\theta}^{(r)})$ by

$$w_k^{(r)} := \frac{G_k^{(r)}}{\sum_{j=1}^{K} G_j^{(r)}}, \text{ where}$$

$$\begin{aligned} G_k^{(r)} &:= \exp(L_k^{(r)} - L_{max}^{(r)}) \\ L_k^{(r)} &:= \log\left(P(\text{data}|\boldsymbol{\theta}_k^{(r)}, M = M_k)P(\boldsymbol{\theta}_k^{(r)}|M = M_k)P(M = M_k)\right) \\ L_{max}^{(r)} &:= \max_{k=1,\cdots,K} L_k^{(r)} \end{aligned}$$

Therefore

$$\frac{1}{R}\sum_{r=1}^{R} w_k^{(r)} \text{ estimates } P(M = M_k|\text{data}).$$

K small: Model selection for financial return data

Since zero in 95% credible interval for $\rho_{MB|T}$ and the posterior mode of $\rho_{ST|M}$ and $\rho_{MB|T}$ small we want to use Congdon (2006)'s method if these copula pairs needed

Model	PCC	Formula	$\widehat{P(M_k\|\text{data})}$
M_1:	with all pairs	$c_{SM}\, c_{MT}\, c_{TB}\, c_{ST\|M}\, c_{MB\|T}\, c_{SB\|MT}$	0.00071
M_2:	without $c_{MB\|T}$	$c_{SM}\, c_{MT}\, c_{TB}\, c_{ST\|M}\, c_{SB\|MT}$	0.25900
M_3:	without $c_{MB\|T}$ and $c_{ST\|M}$	$c_{SM}\, c_{MT}\, c_{TB}\, c_{SB\|MT}$	0.68807
M_4:	without $c_{MB\|T}$,$c_{ST\|M}$ and c_{TB}	$c_{SM}\, c_{MT}\, c_{SB\|MT}$	0.05220

K small: Model selection for financial return data

Many of the posterior estimates of the degree of freedom are large, so want to check if we can use for these pairs a Gaussian copula

Model	PCC	$\widehat{P(M_k \mid \text{data})}$
M_1:	All pairs are $t-$copulas	$1.74 \cdot 10^{-13}$
M_2:	First pair c_{SM} is $t-$copula all other pairs are Gaussian copulas	0.9999924
M_3:	All pairs are Gaussian copulas	$7.62 \cdot 10^{-6}$

Only for c_{SM} a t-copula is needed.

Bayesian Model Selection: K large (RJMCMC)

- If a single pair copula type is used in a PCC of dimension n, then all other possible PCC's of dimension n provide a factorization of the same joint density
- So we want to find reduced PCC's which best fits the data, i.e. let the data discover conditional independence conditions.
- Identifiability: Use only one decomposition
- $n = 3$:

Full PCC's:	$c_{12}c_{23}c_{13\|2}$
Reduced by 1 pair-copula:	$c_{12}c_{23}, c_{12}c_{13\|2}, c_{23}c_{13\|2}$
Reduced by 2 pair copulas:	$c_{12}, c_{23}, c_{13\|2}$

Bayesian framework of PCC's with $n = 3$

- Each of the 7 different PCC's with $n \leq 3$ is identified by a model index $\mathbf{m} = (m_1, m_2, m_3) = (i_1 j_1, i_2 j_2, i_3 j_3 | k_3)$.
  Example: $c_{12} c_{23|1}$ corresponds to $\mathbf{m} = (12, 00, 23|1)$.
- Each PCC model $\mathbf{m}$ has parameter vector $\boldsymbol{\theta}_m = (\boldsymbol{\theta}_{m_1}, \boldsymbol{\theta}_{m_2}, \boldsymbol{\theta}_{m_3})$.
  Example: Model $\mathbf{m} = (12, 23, 00|0)$ has parameter vector $\boldsymbol{\theta}_m = (\boldsymbol{\theta}_{12}, \boldsymbol{\theta}_{23}, \boldsymbol{\theta}_{00|0}) = (\boldsymbol{\theta}_{12}, \boldsymbol{\theta}_{23})$.
- Goal is to estimate the best fitting model $\mathbf{m}$ and the corresponding parameter vector $\boldsymbol{\theta}_m$ using a Bayesian approach, i.e. the model $\mathbf{m}$ and θ_m are considered random quantities.
- Inference about $\mathbf{m}$ and θ_m is done via the joint posterior distribution of $(\mathbf{m}, \boldsymbol{\theta}_m)$.

MCMC algorithm for PCC's with $n = 3$

- Problems:
  - Want to facilitate estimation without having to fit all models
  - Joint posterior is not analytically tractable
- Approach: Construct a Markov Chain Monte Carlo (MCMC) algorithm which simultaneously estimates $\mathbf{m}$ and $\boldsymbol{\theta}_m$.
- Requirement: Need to accommodate varying model dimension, i.e. construct MCMC iterates $\boldsymbol{\theta}_{\mathbf{m}^r} = (\boldsymbol{\theta}^r_{m_1^r}, \boldsymbol{\theta}^r_{m_2^r}, \boldsymbol{\theta}^r_{m_3^r})$ where $\mathbf{m}^r$ is the current model at iteration r.
- Solution: Reversible jump (RJ) MCMC proposed by (Green 1995)

General RJ MCMC (1)

- algorithm stays in current model using a Metropolis Hastings (MH) step
- algorithm moves to a larger model using a MH step (birth)
- algorithm moves to a smaller model using a MH step (death)

General RJ MCMC (2)

Model	M_1		M_2
Parameter	$\boldsymbol{\theta}^{(1)} \in \mathbf{R}^{d_1}$		$\boldsymbol{\theta}^{(2)} \in \mathbf{R}^{d_2}$
	d_1	$<$	d_2
Proposal	$\boldsymbol{\eta}^{(1)} \sim \varphi_1(\cdot)$		$\boldsymbol{\eta}^{(2)} \sim \varphi_2(\cdot)$
	$\boldsymbol{\eta}^{(1)} \in \mathbf{R}^{a_1}$		$\boldsymbol{\eta}^{(2)} \in \mathbf{R}^{a_2}$
Dimension matching	$d_1 + a_1$	$=$	$d_2 + a_2$
Bijection	$\begin{pmatrix} \boldsymbol{\theta}^{(1)} \\ \boldsymbol{\eta}^{(1)} \end{pmatrix}$	$\leftrightarrow$	$\begin{pmatrix} \boldsymbol{\theta}^{(2)} \\ \boldsymbol{\eta}^{(2)} \end{pmatrix}$

Acceptance probability for MH step from M_1 to M_2

$$\alpha(\boldsymbol{\theta}^{(1)}, \boldsymbol{\theta}^{(2)}) = \min\{1, A\}, \text{ where}$$

$$A := \frac{p(2, \boldsymbol{\theta}^{(2)}|\mathbf{y})}{p(1, \boldsymbol{\theta}^{(1)}|\mathbf{y})} \cdot \frac{p_{2\to1}}{p_{1\to2}} \cdot \frac{\varphi_2(\boldsymbol{\eta}^{(2)})}{\varphi_1(\boldsymbol{\eta}^{(1)})} \cdot \left| \frac{\partial\left(\boldsymbol{\theta}^{(2)}, \boldsymbol{\eta}^{(2)}\right)}{\partial\left(\boldsymbol{\theta}^{(1)}, \boldsymbol{\eta}^{(1)}\right)} \right|$$

- $p(k, \boldsymbol{\theta}^{(k)}|y)$ = joint posterior density of M_k and $\boldsymbol{\theta}^{(k)}$
- $p_{i\to j}$ = prior switching probability from M_i to M_j
- $\varphi_k(\boldsymbol{\eta}^{(k)})$ = proposal distribution of $\boldsymbol{\eta}^{(k)}$ (suitably chosen)
- $\left| \frac{\partial\left(\boldsymbol{\theta}^{(2)}, \boldsymbol{\eta}^{(2)}\right)}{\partial\left(\boldsymbol{\theta}^{(1)}, \boldsymbol{\eta}^{(1)}\right)} \right|$ = Jacobian of bijection

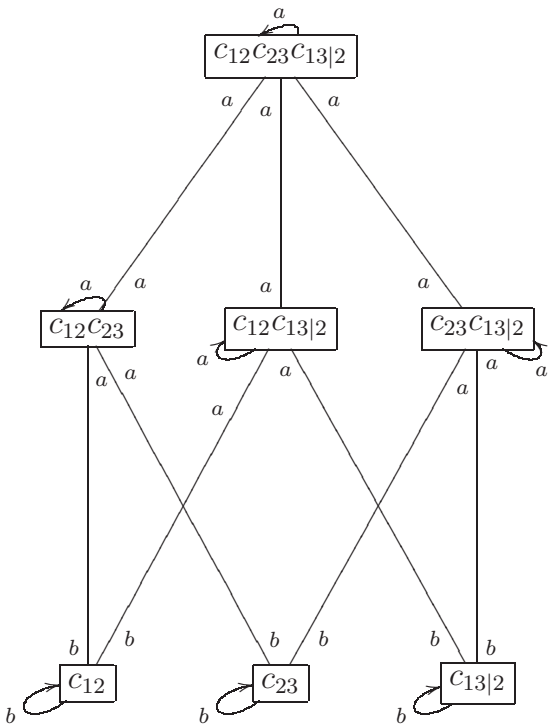
For moves from M_2 to M_1 we use A^{-1}.

RJ MCMC moves for selecting PCC's (n=3)

Notation	Meaning	Example
B0D0	stay move	$c_{12} \longrightarrow c_{12}$
B1D0	birth of 1 factor	$c_{12} \longrightarrow c_{12}\, c_{23}$
D1B0	death of 1 factor	$c_{12}\, c_{23} \longrightarrow c_{12}$

Graph of all PCC's for $n = 3$

$a = \frac{1}{4}$, $b = \frac{1}{3}$



With probabilities a or b the chain moves to other model or stays in the current model

Acceptance probability for B1D0 move

Actual state	$\theta^o := \theta^o_{\mathbf{m}^o}$	$\mathbf{m}^o = (m_1^o, m_2^o, m_3^o)$
Proposed state	$\theta^p := \theta^p_{\mathbf{m}^p}$	$\mathbf{m}^p = (m_1^p, m_2^p, m_3^p)$
birth index	$\{m_s^p\} :=$	$\mathbf{m}^p \backslash \mathbf{m}^o$
stay indices	$\{m_v^o, m_w^o\} :=$	$\mathbf{m}^p \cap \mathbf{m}^o$

Example: $(\theta^p_{12}, \theta^p_{23}) \to (\theta^p_{12}, \theta^p_{23}, \theta^p_{13|2})$

Bijection:

$$\theta^p := \begin{pmatrix} \theta^p_{m_s^p} \\ \theta^p_{m_v^o} \\ \theta^p_{m_w^o} \end{pmatrix} = \begin{pmatrix} \eta^o_{m_s^p} \\ \theta^o_{m_v^o} \\ \theta^o_{m_w^o} \end{pmatrix} =: \begin{pmatrix} \eta^o_{m_s^p} \\ \theta^o \end{pmatrix}$$

Here $\eta^o_{m_s^p}$ is bivariate normal centered at $\theta_{m_s^p, last}$ with specified covariance matrix

MLE for SMT- data

Parameter	*MLE*
ρ_{SM}	-0.25
ρ_{MT}	0.47
$\rho_{ST\|M}$	-0.11
ν_{SM}	4.22
ν_{MT}	16.45
$\nu_{ST\|M}$	300
log.likelih.	200.9

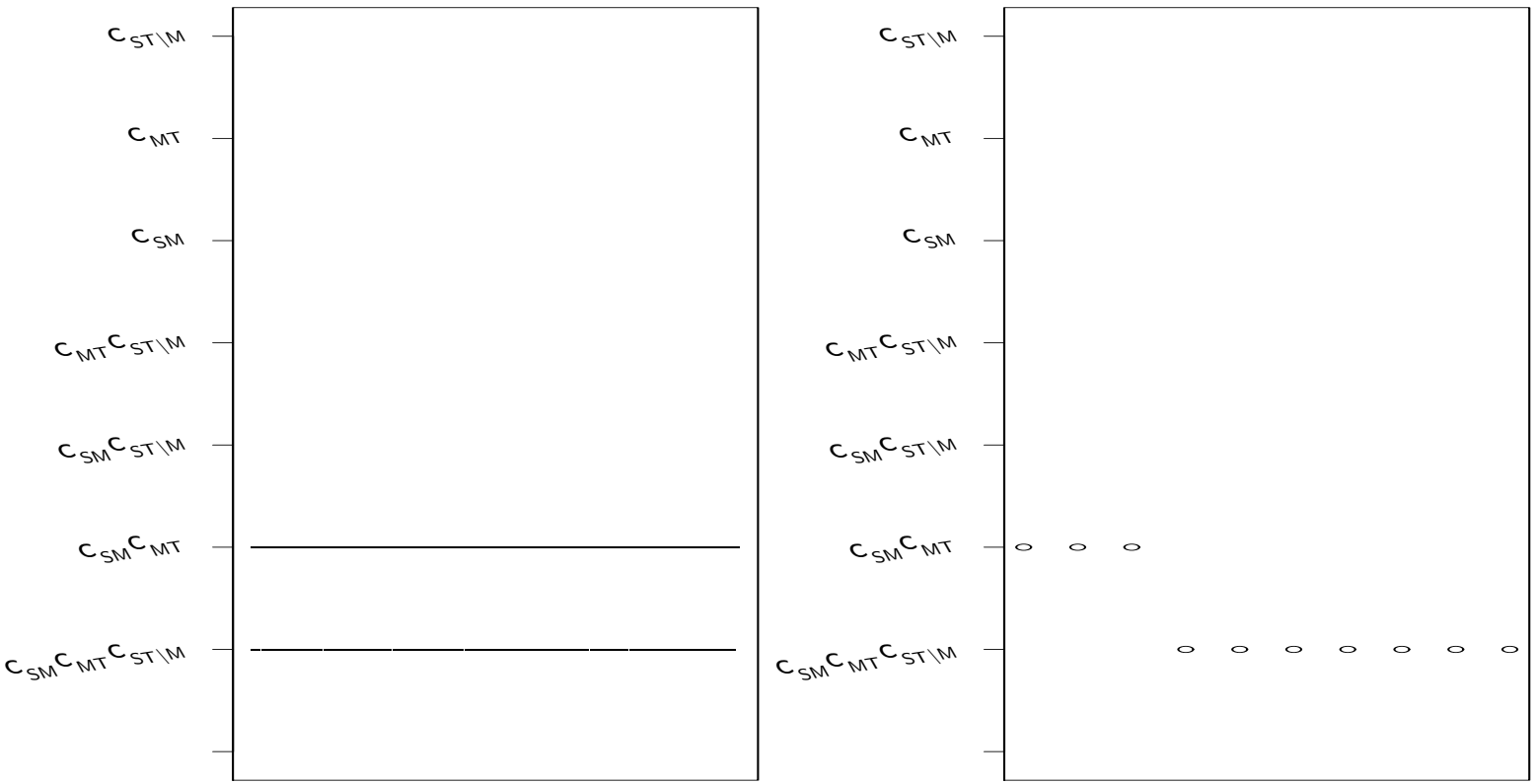
Is the correlation $\rho_{ST|M} = -0.11$ significantly negligible?

Model notation

$c_{SM}\, c_{MT}\, c_{ST\|M}$	M_1
$c_{SM}\, c_{MT}$	M_2
$c_{SM}\, c_{ST\|M}$	M_3
$c_{MT}\, c_{ST\|M}$	M_4
c_{SM}	M_5
c_{MT}	M_6
$c_{ST\|M}$	M_7

Would Model M_2 now be preferred to Model M_1?

Financial returns: Statistics of visited models



Model M_1 was visited 2872 times out of 10000. Model M_2 was visited 7128 times out of 10000.

Summary statistics for visited Models

Model M_2: $c_{SM}\, c_{MT}$

	2.5% Quantile	50% Quantile	97.5% Quantile	Est. Post. Mean	Est. Post. Mode	MLE
ρ_{SM}	-0.31	-0.25	-0.18	-0.25	-0.25	-0.25
ν_{SM}	3.39	4.51	6.86	4.65	4.39	4.21
ρ_{MT}	0.42	0.47	0.50	0.47	0.47	0.47
ν_{MT}	12.64	270.89	944.71	343.05	122.94	16.65

Model M_1: $c_{SM}\, c_{MT}\, c_{ST|M}$

	2.5% Quantile	50% Quantile	97.5% Quantile	Est. Post. Mean	Est. Post. Mode	MLE
ρ_{SM}	-0.32	-0.25	-0.18	-0.25	-0.25	-0.25
ν_{SM}	3.31	4.61	7.08	4.71	4.49	4.24
ρ_{MT}	0.43	0.47	0.52	0.47	0.47	0.47
ν_{MT}	16.24	264.85	861.86	305.85	156.72	16.45
$\rho_{ST\|M}$	-0.17	-0.11	-0.04	-0.11	-0.11	-0.11
$\nu_{ST\|M}$	126.20	498.14	958.40	527.08	420.38	300.00

Posterior model probabilities

RJMCMC

Model	PCC	Formula	$\widehat{P(M_k\|\text{data})}$
M_1	with all pairs	$c_{SM}\,c_{MT}\,c_{ST\|M}$	0.29
M_2	without $c_{ST\|M}$	$c_{SM}\,c_{MT}$	0.71
M_3	without c_{MT}	$c_{SM}\,c_{ST\|M}$	0
M_4	without c_{SM}	$c_{MT}\,c_{ST\|M}$	0
M_5	without c_{MT} and $c_{ST\|M}$	c_{SM}	0
M_6	without c_{SM} and $c_{ST\|M}$	c_{MT}	0
M_7	without c_{SM} and c_{MT}	$c_{ST\|M}$	0

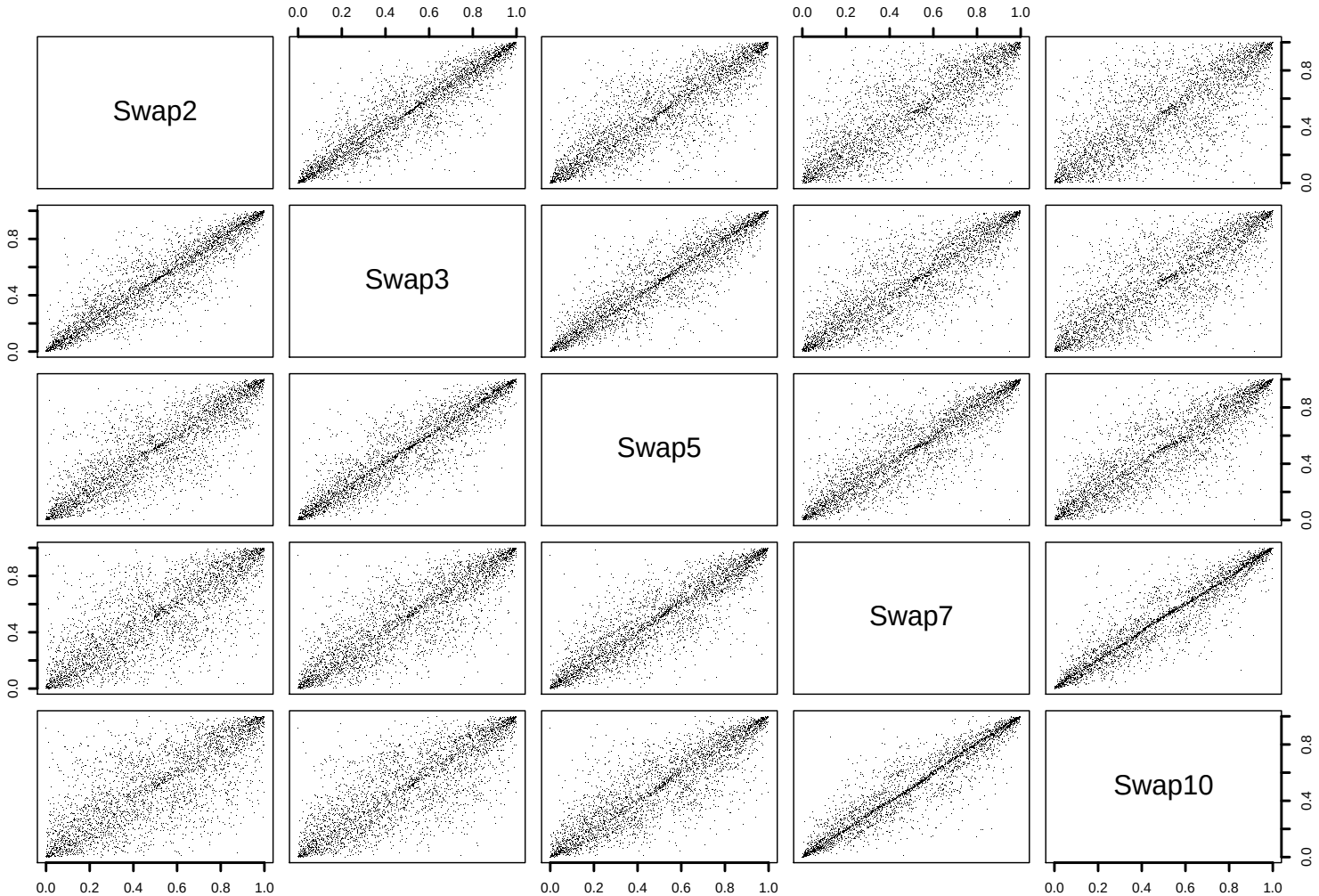
Congdon

Model	PCC	Formula	$\widehat{P(M_k\|\text{data})}$
M_1:	with all pairs	$c_{SM}\,c_{MT}\,c_{ST\|M}$	0.22
M_2:	without $c_{ST\|M}$	$c_{SM}\,c_{MT}$	0.78

Euro Swap Rates

- daily swap rates with $2, 3, 5, 7$ and 10 year maturity quoted in Euro's
- considered period is Dec 7, 1988 until June 21, 2001
- serial marginal correlation removed with an ARMA(1,1)-GARCH(1,1) model and standardized residuals transformed empirically to data with uniform margins
- Notation: Swap2, Swap3, Swap5, Swap7 and Swap10 or abbreviated $\mathbf{S}_1, ..., \mathbf{S}_5$

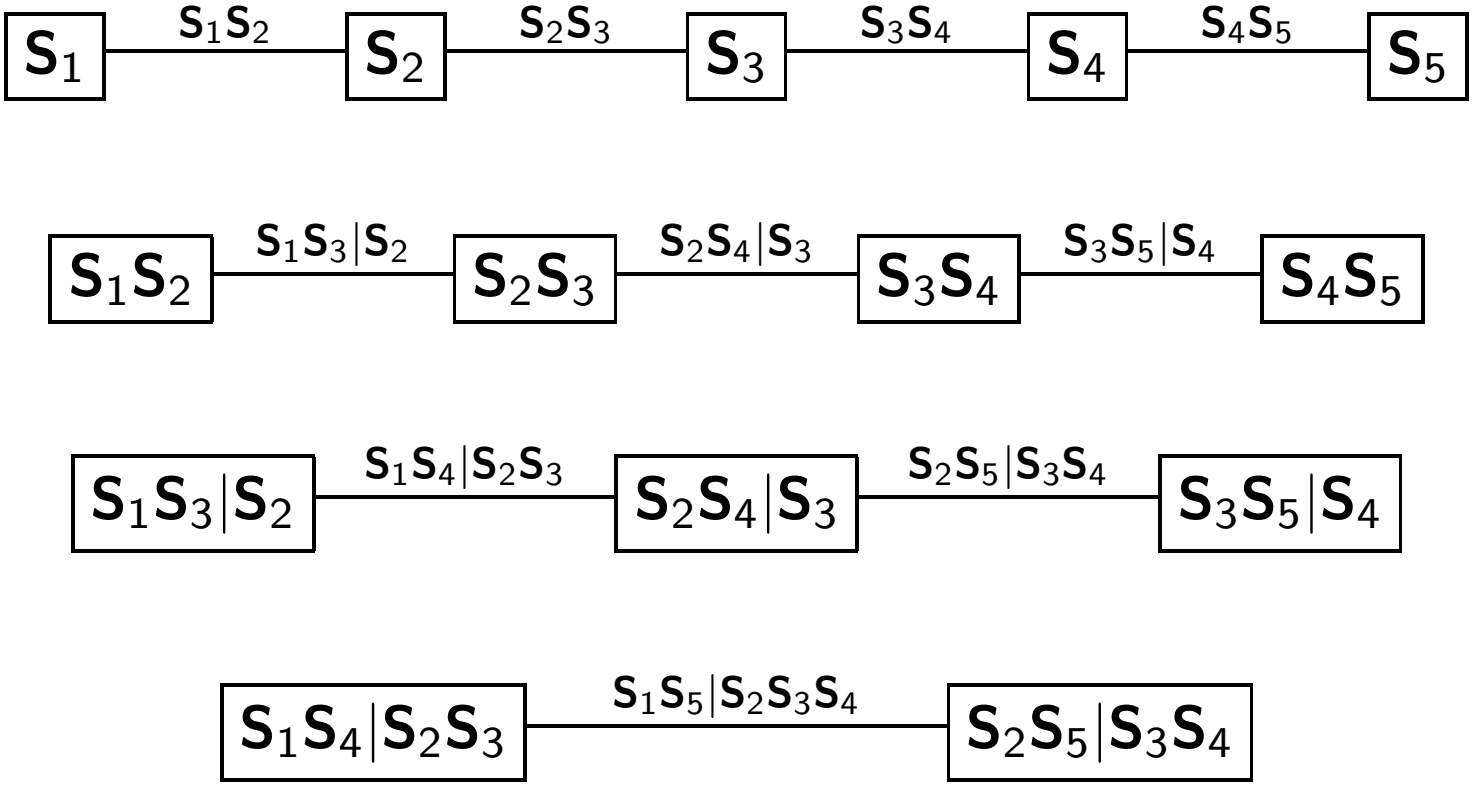
Dependence structure of Euro Swap Rates



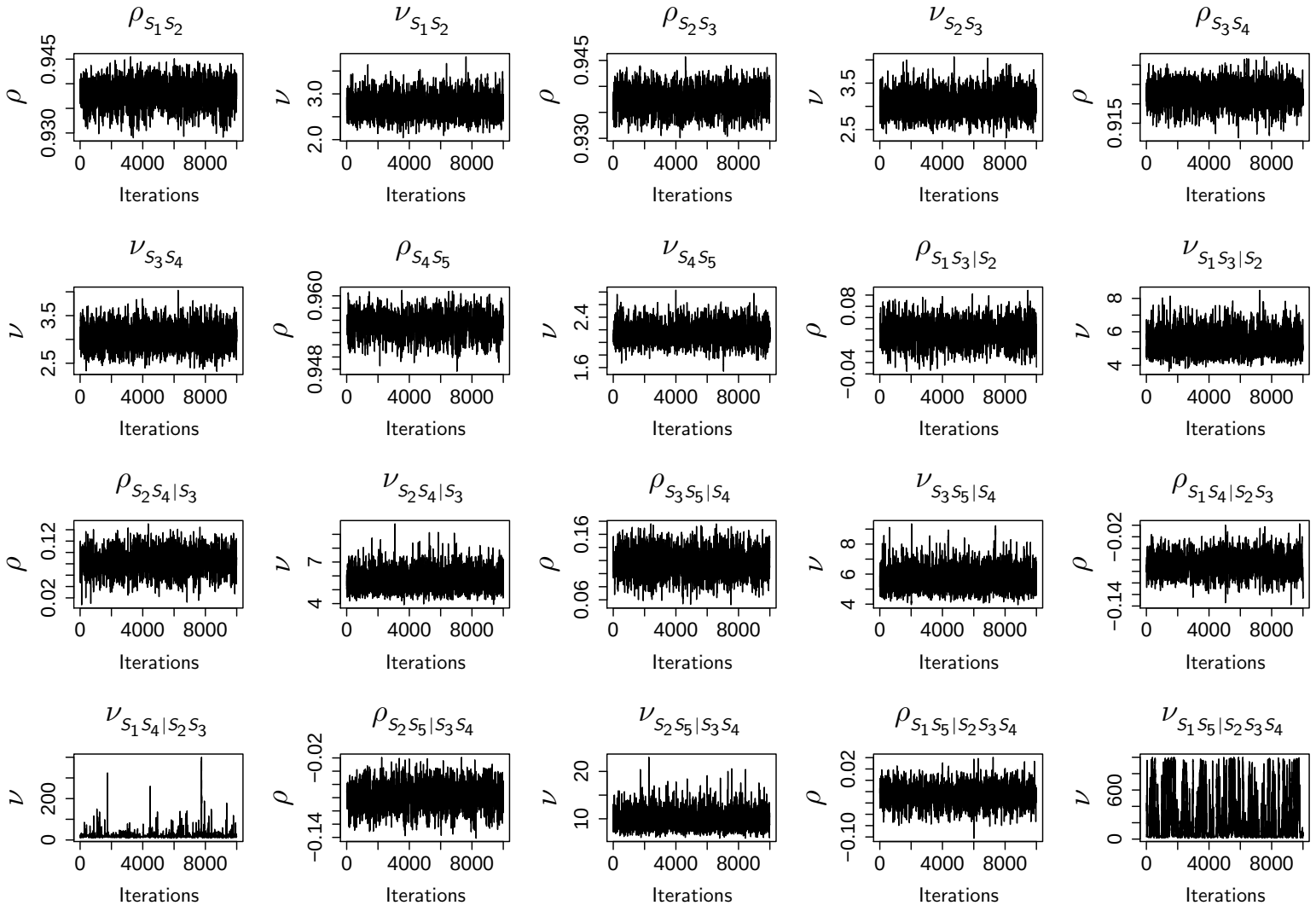
Estimated df's of bivariate t-copula margins for swap rates

	Swap3	**Swap5**	**Swap7**	**Swap10**
Swap2	**2.87**	3.62	4.53	5.23
Swap3		**2.86**	4.19	4.93
Swap5			**3.00**	3.56
Swap7				**2.07**

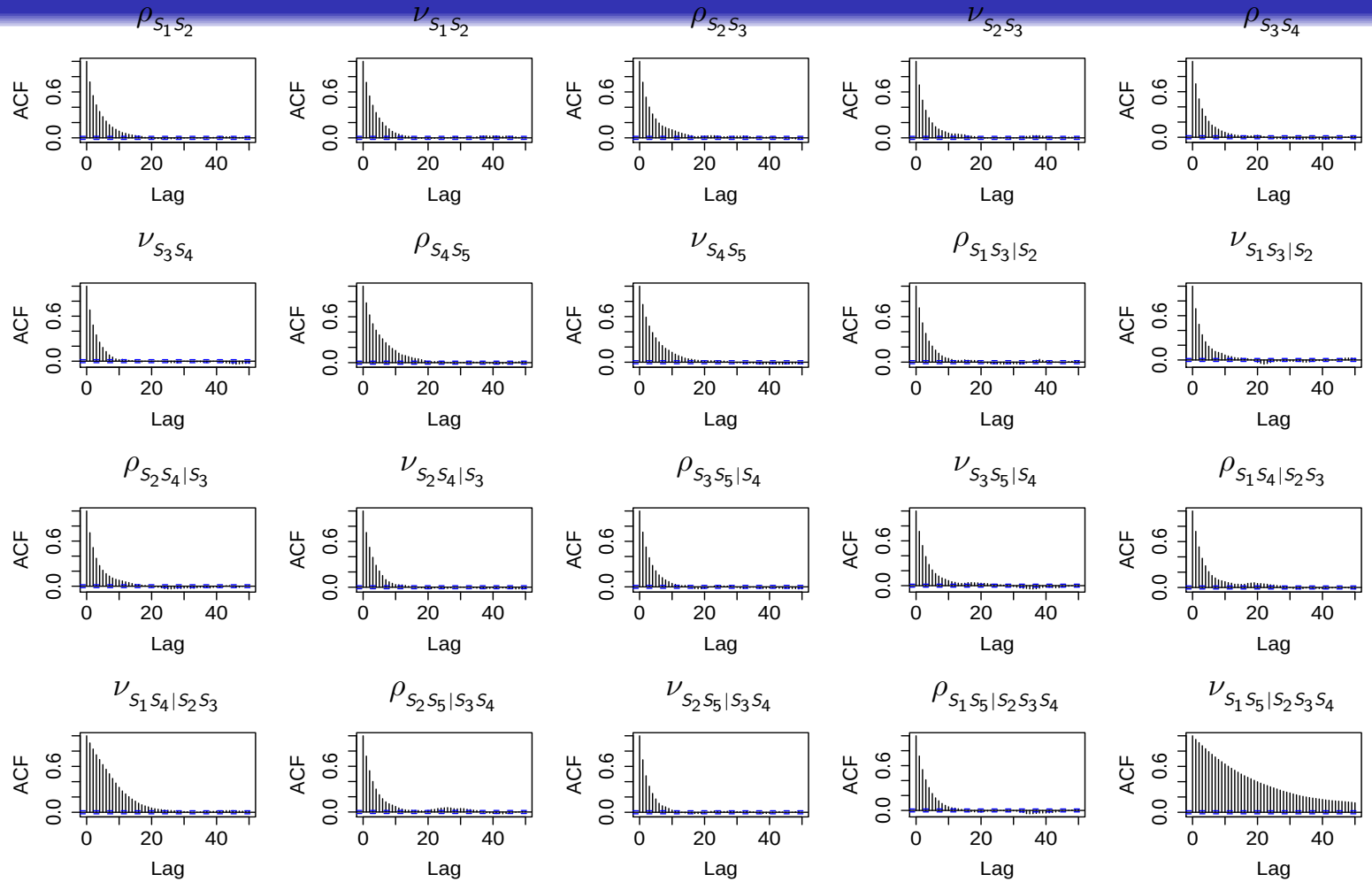
Chosen D-vine structure for swap rates



Trace of MCMC iterations for swap rates

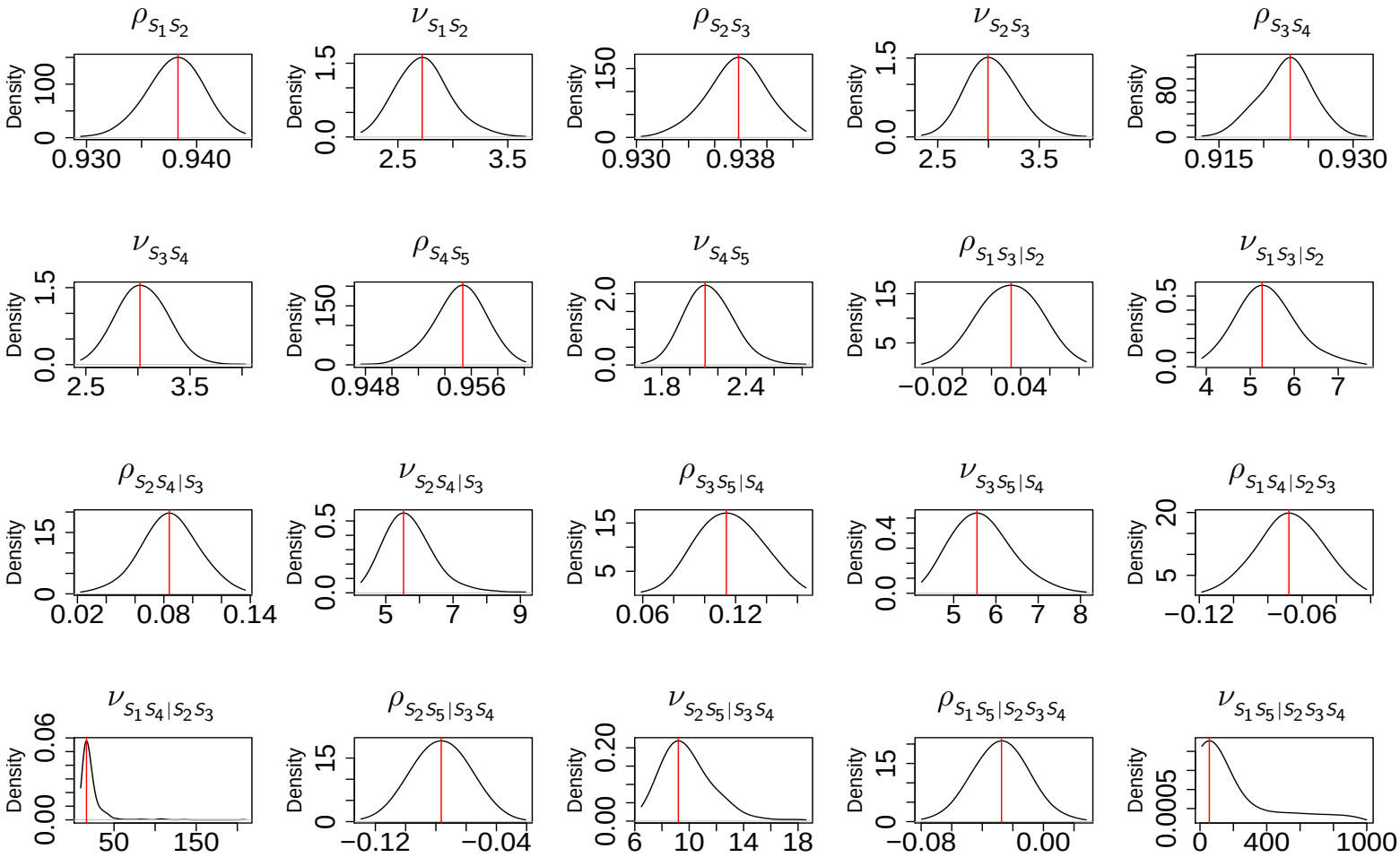


Autocorrelations among the MCMC iterations for the swap rates



For the further analysis we used burnin 1000 iteration and only every 20th iteration.

Estimated posterior densities of D-vine parameters



Summary statistics for thinned MCMC

	2.5% Quantile	50% Quantile	97.5% Quantile	Est. Post. Mean	Est. Post. Mode	MLE
$\rho_{S_1 S_2}$	0.93	0.94	0.94	0.94	0.94	0.94
$\nu_{S_1 S_2}$	2.33	2.73	3.26	2.74	2.72	2.49
$\rho_{S_2 S_3}$	0.93	0.94	0.94	0.94	0.94	0.94
$\nu_{S_2 S_3}$	2.62	3.03	3.54	3.04	3.00	3.08
$\rho_{S_3 S_4}$	0.92	0.92	0.93	0.92	0.92	0.92
$\nu_{S_3 S_4}$	2.62	3.03	3.48	3.04	3.02	2.87
$\rho_{S_4 S_5}$	0.95	0.96	0.96	0.96	0.96	0.96
$\nu_{S_4 S_5}$	1.86	2.12	2.50	2.14	2.11	2.27
$\rho_{S_1 S_3 \mid S_2}$	-0.01	0.03	0.07	0.03	0.03	0.03
$\nu_{S_1 S_3 \mid S_2}$	4.19	5.32	7.00	5.40	5.27	5.08
$\rho_{S_2 S_4 \mid S_3}$	0.04	0.08	0.12	0.08	0.08	0.08
$\nu_{S_2 S_4 \mid S_3}$	4.63	5.63	7.36	5.70	5.53	5.30
$\rho_{S_3 S_5 \mid S_4}$	0.08	0.12	0.15	0.12	0.11	0.12
$\nu_{S_3 S_5 \mid S_4}$	4.64	5.62	7.22	5.67	5.55	5.58
$\rho_{S_1 S_4 \mid S_2 S_3}$	-0.10	-0.07	-0.03	-0.07	-0.07	-0.07
$\nu_{S_1 S_4 \mid S_2 S_3}$	10.84	18.12	55.27	22.21	16.48	11.94
$\rho_{S_2 S_5 \mid S_3 S_4}$	-0.11	-0.08	-0.04	-0.08	-0.08	-0.08
$\nu_{S_2 S_5 \mid S_3 S_4}$	7.10	9.51	14.58	9.90	9.20	7.69
$\rho_{S_1 S_5 \mid S_2 S_3 S_4}$	-0.06	-0.03	0.01	-0.03	-0.03	-0.02
$\nu_{S_1 S_5 \mid S_2 S_3 S_4}$	15.37	64.51	954.21	223.08	55.84	16.30

Log likelihoods

Parameter values	**loglikelihood**
multivariate t with common df	13937.15
MLE	14036.83
Posterior Mode	14037.96

- Likelihood Ratio Test rejects a multivariate t copula with common df decisively, estimated common df is 3.36.
- Since posterior mode log likelihood is close to MLE likelihood, prior choice for ρ and ν are close to being noninformative in this data set.

Model selection for swap rate data

Model	PCC	$P(\widehat{M_k \mid \text{data}})$
M_1:	with all pairs	0.19
M_2:	without $c_{S_1 S_5 \mid S_2 S_3 S_4}$	0.81
M_3:	without $c_{S_1 S_5 \mid S_2 S_3 S_4}$ and $c_{S_2 S_5 \mid S_3 S_4}$	2.5×10^{-8}
M_4:	without $c_{S_1 S_5 \mid S_2 S_3 S_4}$, $c_{S_2 S_5 \mid S_3 S_4}$ and $c_{S_1 S_4 \mid S_2 S_3}$	5.2×10^{-11}

Summary and Conclusion

- PCC's such as canonical and D-Vines allow for very flexible class of multivariate distributions.
- The Bayesian approach can solve estimation as well as model selection problems. It also gives credible intervals for parameters of interest.
- The proposed RJMCMC Bayesian algorithm for $n > 3$ is under implementation. Incorporation of pair-copulas from different parametric families is under consideration.
- The proposed RJMCMC algorithm is an alternative to DIC and other Bayesian model choice algorithms
- Congdon's method can be used to make model comparisons for a small number of models

Aas, K., C. Czado, A. Frigessi, and H. Bakken (2007).
Pair-copula constructions of multiple dependence.
Insurance, Mathematics and Economics, DOI 10.1016/j.insmatheco.2007.02.001, preprint available at SFB 386 Statistical Analysis of Diskrete Structures, www.stat.uni–muenchen.de/sfb386/.

Bedford, T. and R. M. Cooke (2001).
Probability density decomposition for conditionally dependent random variables modeled by vines.
Annals of Mathematics and Artificial Intelligence 32, 245–268.

Bedford, T. and R. M. Cooke (2002).
Vines - a new graphical model for dependent random variables.
Annals of Statistics 30(4), 1031–1068.

Berg, D. and K. Aas (2007).
Models for construction of higher-dimensional dependence.

Congdon, P. (2006).
Bayesian model choice based on monte carlo estimates of posterior model probabilities.
Computational Statistics & Data Analysis 50, 346–357.

Green, P. (1995).
Reversible jump markov chain monte carlo computation and Bayesian model determination.
Biometrika 82, 711–732.

Joe, H. (1996).
Families of m-variate distributions with given margins and m(m-1)/2 bivariate dependence parameters.
In L. Rüschendorf and B. Schweizer and M. D. Taylor (Ed.), *Distributions with Fixed Marginals and Related Topics*.

Kurowicka, D. and R. Cooke (2006).
Uncertainty analysis with high dimensional dependence modelling.

Chichester: Wiley.

Kurowicka, D. and R. M. Cooke (2004).
Distribution - free continuous bayesian belief nets.
In *Fourth International Conference on Mathematical Methods in Reliability Methodology and Practice, Santa Fe, New Mexico*.

Scott, S. L. (2002).
Bayesian methods for hidden Markov models: Recursive computing in the 21st century.
J. Amer. Statist. Assoc. 97, 337–351.