Technische Universität München

Zentrum Mathematik

Joint estimation of parameters in multivariate normal regression with correlated errors using pair-copula constructions and an application to finance

Diplomarbeit

von

Jasper Niklas Lanzendörfer

Themenstellerin:	Prof. Dr. Claudia Czado
Betreuer:	Prof. Dr. Claudia Czado
	Dr. Aleksey Min
Abgabetermin:	27. Juli 2009

Hiermit erkläre ich, dass ich die Diplomarbeit selbstständig angefertigt und nur die angegebenen Quellen verwendet habe.

Garching, den 27. Juli 2009

Acknowledgments

First of all, I would like to express my sincere gratitude to Prof. Dr. Claudia Czado for making this diploma thesis possible and for the excellent, intensive supervision throughout the whole developing process. This thesis has gained a lot from her numerous hints and from the valuable discussions during our frequent meetings.

Furthermore, I owe special thanks to Dr. Aleksey Min who always had time to answer my often tricky questions and helped me to solve my encountered problems with his expert guidance.

Dr. Carlos Almeida supported me during the last phase of my research and Prof. Michael Smith helped me to gain the data I analyzed. I gratefully acknowledge here my sincere thanks to both of them.

Last but not least, I would like to thank my parents who unremittingly supported me during my years of study and my wife Lupe for her encouragement, understanding, patience and love during my whole studies and particularly during the difficult periods of this work.

Contents

List of Figures

List of Tables

List of Algorithms

Chapter 1

Introduction

In finance, the capital asset pricing model (CAPM) is frequently used to describe the relationship between the returns of a well-diversified portfolio and the market return. The CAPM may be viewed as multivariate normal regression model with a common regression variable – the market return.

The aim of this thesis is to provide an MCMC procedure for the joint estimation of regression parameters, residual variance and residual correlation in such a model environment. An important part of this is to model the residual correlation with a Gaussian pair-copula construction instead of a correlation or precision matrix. This approach allows for flexibility and model reduction.

The thesis is organized as follows: In Chapter 2, we recall some basics that we later need in the thesis, like Bayesian inference, Markov Chain Monte Carlo methods and pair-copula constructions.

In Chapter 3, we define the model for two dimensions and develop a fast running algorithm that estimates the model parameters. Finally, we verify the well behavior of the algorithm by performing a small sample analysis which comprises of a wide number of scenarios.

In Chapter 4, we consider the general case of an arbitrary dimension. Again, our model is defined and an MCMC algorithm is developed. For the estimation of dependence structure covered by a correlation matrix, we will propose two different ways of performing the Metropolis-Hastings step.

In Chapter 5, we look at a data set of U.S. industrial returns covering a period of over 80 years. After an extensive data analysis, we choose 3 suitable structures and apply our previously derived MCMC algorithm on the data. The results of the MCMC are analyzed, compared and searched for possible model reduction. The chapter concludes with a validation of the results.

In Chapter 6, we summarize the results of the thesis and propose sources for possible extensions.

Chapter 2

Preliminaries

2.1 Moments and their estimation

In this section we briefly recall the definitions and meanings of moments and how they can be estimated if one has a sample of i.i.d. random variables.

Definition Let $X, X_1, \dots X_n$ be i.i.d. with distribution function F. Then the *kth moment of X* is defined by $\mu_k := E(X^k) = \int X^k \mathrm{d}F$ if the integral exists, otherwise the kth moment does not exist. If it exists, the kth moment can be estimated by $m_k := \frac{1}{n}\sum_{i=1}^{n} X_n^k$.

The first moment is called the expectation or mean of X, its estimator m_1 is called the mean of $X_1, \dots, X_n$ and also denoted by $\bar{X}$. If μ_k exists for all $k \in \mathbb{N}$, then the distribution of X is uniquely determined by the set of all moments $\{\mu_k, k \in \mathbb{N}\}$.

The value of the moments μ_k with large $k \in \mathbb{N}$ often depends strongly on the expectation $E(X)$. Thus, one often prefers another measure, the centralized moment, which is defined as follows:

Definition Let $X, X_1, \dots X_n$ be i.i.d. with distribution function F. If the kth moment exists, the *kth centralized moment of X* is defined by

$$\widetilde{\mu}_k := E((X - \mu_1)^k) = \int (X - E(X))^k \mathrm{d}F.$$

If it exists, the kth centralized moment can be estimated by $m_k := \frac{1}{n}\sum_{i=1}^{n}(X_n - \bar{X})^k$.

The first centralized moment is always 0, the second centralized moment is called *variance* and its root the *standard deviation*. Other important moment measures are the *skewness* and the *kurtosis*. Definitions of mean, variance, skewness and kurtosis and corresponding estimators are provided in table 2.1.

	Notation	Definition	Estimator
Mean	μ	$E(X)$	$\frac{1}{n}\sum_{i=1}^{n} X_i$
Variance	σ^2	$E((X-\mu^2)$	$\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X})^2$
Skewness	γ	$\frac{E((X-E(X))^3)}{\sigma^3}$	$\frac{\frac{1}{n}\sum_{i=1}^{n}(X_i-\bar{X})^4}{(\sum_{i=1}^{n}(X_i-\bar{X})^2)^{\frac{3}{2}}}$
(Excess) Kurtosis	κ	$\frac{E((X-\mu)^4)}{\sigma^4}-3$	$\frac{\frac{1}{n}\sum_{i=1}^{n}(X_i-\bar{X})^4}{(\sum_{i=1}^{n}(X_i-\bar{X})^2)^2}-3$

Table 2.1: Definition of mean, variance, skewness and kurtosis as well as their estimators

The mean of a distribution tells us what we may expect as typical value for X, or, following the law of large numbers, what the mean of a large sample will approximately be. The variance measures the typical quadratic distance between a single value drawn from the distribution of X and the mean. To get information about the lopsidedness of the distribution, one can look at the skewness. A symmetric distribution like the normal distribution has a skewness of 0, whereas a negative skewness shows that the distribution is skewed to the left, a positive skewness shows the opposite. At last, the kurtosis is a measure for the probability of extreme events. While the normal distribution has a kurtosis of zero, a distribution with "fatter tails" has a positive kurtosis, whereas a negative kurtosis shows that extreme events are even less likely as in the normal case. From its definition it follows that the kurtosis is always greater than -3, except for point distributions, for which the kurtosis is equal to -3. In the literature one also finds a definition of kurtosis where the "-3"-Term is omitted; this is why our definition is often referred to as "excess kurtosis".

2.2 Normal distribution

In this section we recall the definition and some basic properties of the normal distribution, which we will need later in this thesis. They can be found for instance in Bickel and Doksum (2001).

Definition *Multivariate normal distribution.* Let $\boldsymbol{\mu} \in \mathbb{R}^d$ and let $\Sigma \in \mathbb{R}^{d\times d}$ be a positive definite matrix. A random vector $\boldsymbol{X}$ has a d dimensional normal distribution with mean vector $\boldsymbol{\mu}$ and covariance matrix Σ, Notation $\boldsymbol{X} \sim \mathcal{N}_d(\mu, \Sigma)$, if its density can be written as

$$p(\boldsymbol{x}) = \frac{1}{\sqrt{(2\pi)^n \det(\Sigma)}} \exp\left\{-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})'\Sigma^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\right\}$$

where $\boldsymbol{x} = (x_1, \ldots, x_d)'$.

Example *Bivariate normal distribution.* Let $\boldsymbol{\mu} \in \mathbb{R}^2$, $\boldsymbol{\sigma} \in (0,\infty)^2$ and $\rho \in (-1,1)$. The density of $\boldsymbol{X} \sim \mathcal{N}_2\left(\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix} \right)$ is

$$p(x_1,x_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left\{ -\frac{1}{2(1-\rho^2)} \left(\frac{(x_1-\mu_1)^2}{\sigma_1^2} + \frac{(x_2-\mu_2)^2}{\sigma_2^2} - 2\rho\frac{(x_1-\mu_1)(x_2-\mu_2)}{\sigma_1\sigma_2} \right) \right\} \tag{2.1}$$

Remark Let $X \sim \mathcal{N}_d(\boldsymbol{\mu}, \Sigma)$. Let $\boldsymbol{b} \in \mathbb{R}^k$ and $A \in \mathbb{R}^{k\times d}$.
Then it holds

$$\boldsymbol{Z} := A\boldsymbol{X} + \boldsymbol{b} \sim \mathcal{N}_k(A\boldsymbol{\mu} + \boldsymbol{b}, A\Sigma A'). \tag{2.2}$$

Theorem Let $(\boldsymbol{X}', \boldsymbol{Y}')' \sim \mathcal{N}_d(\boldsymbol{\mu}, \Sigma)$ with $d = d_x + d_y$, $\boldsymbol{X} \in \mathbb{R}^{d_x}$, $\boldsymbol{Y} \in \mathbb{R}^{d_y}$, where the mean $\boldsymbol{\mu}$ and the covariance matrix Σ are suitably partitioned as

$$\boldsymbol{\mu} = \begin{pmatrix} \boldsymbol{\mu}_x \\ \boldsymbol{\mu}_y \end{pmatrix} \text{ and } \Sigma = \begin{pmatrix} \Sigma_{xx} & \Sigma_{xy} \\ \Sigma_{xy}' & \Sigma_{yy} \end{pmatrix}$$

with $\boldsymbol{\mu}_x \in \mathbb{R}^{d_x}$, $\boldsymbol{\mu}_y \in \mathbb{R}^{d_y}$, $\Sigma_{xx} \in \mathbb{R}^{d_x\times d_x}$, $\Sigma_{yy} \in \mathbb{R}^{d_y\times d_y}$ and $\Sigma_{xy} \in \mathbb{R}^{d_x\times d_y}$.

Then the random vectors $\boldsymbol{X}$ and $\boldsymbol{Y}$ follow normal distributions, where $\boldsymbol{X} \sim \mathcal{N}_{d_x}(\boldsymbol{\mu}_x, \Sigma_{xx})$ and $\boldsymbol{Y} \sim \mathcal{N}_{d_y}(\boldsymbol{\mu}_y, \Sigma_{yy})$. In addition, also the conditional distribution of $\boldsymbol{X}$ given $\boldsymbol{Y} = \boldsymbol{y}$ is multivariate normal, more precisely $\boldsymbol{X}|\boldsymbol{Y} = \boldsymbol{y} \sim \mathcal{N}_{d_x}(\boldsymbol{\mu}_{x|y}, \Sigma_{x|y})$ with

$$\boldsymbol{\mu}_{x|y} = \boldsymbol{\mu}_x + \Sigma_{xy}\Sigma_{yy}^{-1}(\boldsymbol{y} - \boldsymbol{\mu}_y) \text{ and } \Sigma_{x|y} = \Sigma_{xx} - \Sigma_{xy}\Sigma_{yy}^{-1}\Sigma_{xy}'. \tag{2.3}$$

The conditional covariance matrix $\Sigma_{x|y}$ is the *Schur complement* of Σ_{yy} in the joint covariance matrix Σ.

If we compare the marginal distribution of $\boldsymbol{X}$ and the distribution of $\boldsymbol{X}$ conditional on $\boldsymbol{Y} = \boldsymbol{y}$, we see that the unconditional mean $\boldsymbol{\mu}_x$ is "corrected" by a term that depends on the deviation of the known value $\boldsymbol{y}$ from its mean and on the covariances of $\boldsymbol{X}$ and $\boldsymbol{Y}$. Also the covariance matrix of $\boldsymbol{X}$ changes when the value of $\boldsymbol{Y}$ is known, but this change is independent of the realized value $\boldsymbol{y}$.

2.3 Other distributions and their properties

Besides the normal distribution, we will need some other distributions later in this thesis. Especially, we need distributions for random variables that are constrained to positive

values or values inside a compact interval. One of the simplest is the uniform distribution, which is used for example for non-informative prior distributions on a bounded interval.

Definition *Uniform distribution.* A random variable X has a uniform distribution on the interval $I = (i_1, i_2) \quad (i_1, i_2 \in \mathbb{R}, i_1 < i_2)$, Notation $X \sim Uniform(i_1, i_2)$, if its density can be written as

$$p(x) = \frac{1}{i_2 - i_1}$$

defined over the support (i_1, i_2).

Remark *Properties of the uniform distribution.* Let $X \sim Uniform(i_1, i_2)$.
Then there exists no unique mode for X (since the density is the same on the whole interval (i_1, i_2)), the expectation is $E(X) = \frac{i_1+i_2}{2}$ and the variance is $Var(X) = \frac{(i_2-i_1)^2}{12}$.

Another important distribution is the Gamma distribution for positive random variables, as well as its inverse, the so called Inverse Gamma distribution. There exist many different parametrizations for those distributions in the literature. In our thesis we choose parametrizations which make it very easy to apply inverse transformations.

Definition *Gamma distribution.* A random variable X has a Gamma distribution with parameters $a > 0$ and $b > 0$, Notation $X \sim Gamma(a, b)$, if its density can be written as

$$p(x) = \frac{b^a}{\Gamma(a)} x^{a-1} \exp\{-bx\}$$

defined over the support $(0, \infty)$, where $\Gamma(x) = \int_0^\infty t^{x-1} \exp\{-t\} \mathrm{d}t$ is the Gamma function.

Remark *Properties of the Gamma Distribution.* Let $X \sim Gamma(a, b)$.
Then the mode of X is $\frac{a-1}{b}$ if $a > 1$, the expectation is $E(X) = \frac{a}{b}$ and the variance is $Var(X) = \frac{a}{b^2}$.

Definition *Inverse Gamma distribution.* A random variable X has an Inverse Gamma distribution with parameters $a > 0$ and $b > 0$, Notation $X \sim InverseGamma(a, b)$, if its density can be written as

$$p(x) = \frac{b^a}{\Gamma(a)} x^{-a-1} \exp\left\{-\frac{b}{x}\right\}$$

defined over the support $(0, \infty)$.

Remark *Properties of the Inverse Gamma distribution.* Let $X \sim InverseGamma(a, b)$.
Then the mode of X is $\frac{b}{a+1}$, the expectation is $E(X) = \frac{b}{a-1}$ if $a > 1$, and the variance is $Var(X) = \frac{b^2}{(a-1)^2(a-2)}$ if $a > 2$.

As mentioned before, the relationship between the Gamma distribution and its inverse is very simple.

Corollary Let $X \sim Gamma(a, b)$. Then $X^{-1} \sim InverseGamma(a, b)$.

Besides the uniform distribution, there is another frequently used distribution for continuous random variables defined on bounded intervals, the Beta distribution. Usually one finds this distribution defined on the support (0,1), as follows:

Definition *Beta distribution.* A random variable X has a Beta distribution with parameters $a > 0$ and $b > 0$, Notation $X \sim Beta(a, b)$, if its density can be written as

$$p(x) = \frac{1}{B(a,b)} x^{a-1}(1-x)^{b-1}$$

defined over the support $(0, 1)$, where $B(a, b)$ is the so-called Beta function defined by $B(a, b) := \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$.

Remark *Properties of the Beta distribution.* Let $X \sim Beta(a, b)$.
Then the mode of X is $\frac{a-1}{a+b-2}$ if $a > 1$ and $b > 1$, its expectation is $E(X) = \frac{a}{a+b}$ and the variance is $Var(X) = \frac{ab}{(a+b)^2(a+b+1)}$ if $a > 2$.
If $a = b = 1$, then $X \sim Uniform(0, 1)$.

There is also an important connection between the Gamma distribution and the Beta distribution:

Remark Let X and Y be independent random variables with $X \sim Gamma(a_1, b)$ and $Y \sim Gamma(a_2, b)$. Then $Z := \frac{X}{X+Y} \sim Beta(a_1, a_2)$.

Now one can derive the Beta distribution on an arbitrary interval (i_1, i_2) with $i_1 < i_2 \in \mathbb{R}$ and $i_1 < i_2$ by performing a linear transformation. If $X \sim Beta(a, b)$, we can define $Y := (i_2 - i_1)X + i_1$ and know from the univariate density transformation theorem that the density of Y is obtained by

$$p_Y(y) = \frac{1}{i_2 - i_1} p_X\left(\frac{y - i_1}{i_2 - i_1}\right),$$

if $p_X(\cdot)$ denotes the density of X. We will need later the Beta distribution on $(-1, 1)$, so setting $(i_1, i_2) := (-1, 1)$ leads to

$$p_Y(y) = \frac{1}{2}\frac{1}{B(a,b)}\left(\frac{y+1}{2}\right)^{a-1}\left(\frac{y-1}{2}\right)^{b-1}$$

For $a = b$, it holds

$$p_Y(y) = \frac{2^{-2a+1}}{B(a,a)}(1-y^2)^{a-1}$$

2.4 Bayesian inference

Assume you have data resulting from a statistical model $(\Omega, \mathcal{F}, \{P_{\boldsymbol{\theta}} | \boldsymbol{\theta} \in \boldsymbol{\Theta}\})$ with unknown parameter (vector) $\boldsymbol{\theta}$ located in a parameter space $\boldsymbol{\Theta}$. In the classical approach, one interprets $\boldsymbol{\theta}$ as fixed, but unknown quantity. The drawback of this point of view is that probabilistic interpretations of all functions $g(\boldsymbol{\theta})$ are not possible, so one is limited to interpreting functions of the data like estimators, confidence intervals etc. for fixed values of $\boldsymbol{\theta}$.

An alternative approach is to treat the parameter $\boldsymbol{\theta}$ as a random variable, whose assumed distribution is called the *prior distribution*. This distribution reflects all information available about $\boldsymbol{\theta}$ before observed data is analyzed. If one has a sample $\boldsymbol{x} = (x_1, \ldots, x_n)'$ of i.i.d. random variables $X_1, \ldots, X_n$ with distribution function $F(\cdot|\boldsymbol{\theta})$ for each $\boldsymbol{\theta} \in \boldsymbol{\Theta}$, the additional knowledge about $\boldsymbol{\theta}$ from $\boldsymbol{x}$ is included by calculating the distribution of $\boldsymbol{\theta}$ conditional on $\boldsymbol{x}$, which is called the *posterior distribution*.

To obtain the density $p(\boldsymbol{\theta}|\boldsymbol{x})$ of the posterior distribution, one can apply the well-known theorem of Bayes, if the prior density $p(\boldsymbol{\theta})$ and the likelihood $f(\boldsymbol{x}|\boldsymbol{\theta})$ are available:

Theorem *Bayes theorem*

$$p(\boldsymbol{\theta}|\boldsymbol{x}) = \frac{f(\boldsymbol{x}|\boldsymbol{\theta})p(\boldsymbol{\theta})}{f(\boldsymbol{x})}$$

where $f(\boldsymbol{x}) = \int f(\boldsymbol{x}|\boldsymbol{\theta})p(\boldsymbol{\theta})\,\mathrm{d}\boldsymbol{\theta}$.

Due to the frequent application of this theorem, the presented approach is called *Bayesian inference*. Since $f(\boldsymbol{x})$ does not depend on $\boldsymbol{\theta}$, we usually do not have to calculate it. This means that all available information about $\boldsymbol{\theta}$ is restricted to the prior density and the likelihood, and that the posterior density may be determined by

$$p(\boldsymbol{\theta}|\boldsymbol{x}) \propto f(\boldsymbol{x}|\boldsymbol{\theta})p(\boldsymbol{\theta}) \tag{2.4}$$

up to a proportional constant which guarantees that the integral of the posterior density $p(\boldsymbol{\theta}|\boldsymbol{x})$ is 1.

2.5 Markov Chain Monte Carlo (MCMC) methods

Markov Chain Monte Carlo (MCMC) methods can be used when the distribution of a random variable is hard to calculate. As basis of our presentation of the topic we use the good introductions provided by Gilks et al. (1996) and Gschößl (2006).

Our motivation into the topic is to consider the posterior distribution. In Bayesgian inference, one is often interested e. g. in the posterior mean, mode, variance, quantiles or other characteristics of the posterior distribution.

More generally, if one has data $\boldsymbol{x} \in \mathbb{R}^n$ and a model with unknown parameter vector $\boldsymbol{\theta}$, one wants to know the expected value $E\left(g(\boldsymbol{\theta})|\boldsymbol{x}\right)$ of a function $g(\boldsymbol{\theta})$ that depends on the parameter $\boldsymbol{\theta}$ with respect to the posterior distribution. As we have seen before, the posterior distribution is calculated by $p(\boldsymbol{\theta}|\boldsymbol{x}) = \frac{f(\boldsymbol{x}|\boldsymbol{\theta})p(\boldsymbol{\theta})}{\int f(\boldsymbol{x}|\boldsymbol{\theta})p(\boldsymbol{\theta})\mathrm{d}\boldsymbol{\theta}}$, so the expression of interest can be obtained by

$$E\left(g(\boldsymbol{\theta})|\boldsymbol{x}\right) = \int g(\boldsymbol{\theta})p(\boldsymbol{\theta}|\boldsymbol{x})\mathrm{d}\boldsymbol{\theta} = \frac{\int g(\boldsymbol{\theta})p(\boldsymbol{\theta})f(\boldsymbol{x}|\boldsymbol{\theta})\mathrm{d}\boldsymbol{\theta}}{\int f(\boldsymbol{x}|\boldsymbol{\theta})p(\boldsymbol{\theta})\mathrm{d}\boldsymbol{\theta}}$$

However, in many complex applications it is not possible to calculate the denominator $\int f(\boldsymbol{x}|\boldsymbol{\theta})p(\boldsymbol{\theta})\mathrm{d}\boldsymbol{\theta}$ analytically or in a numerically tractable way, especially for high-dimensional $\boldsymbol{\theta}$. The idea of MCMC methods is to construct a Markov chain with a stationary distribution equal to the posterior distribution. With a sample $\{\boldsymbol{\theta}^{(1)}, \ldots, \boldsymbol{\theta}^{(m)},\ m \in \mathbb{N}\}$ from this Markov chain and the use of the *Law of large numbers for Markov chains*, one can approximate $E\left(g(\boldsymbol{\theta})|\boldsymbol{x}\right)$ by

$$E\left(g(\boldsymbol{\theta})|\boldsymbol{x}\right) \approx \frac{1}{m - r_0} \sum_{r=r_0+1}^{m} g(\boldsymbol{\theta}^{(r)})$$

where $0 \leq r_0 < m$ is the number of samples that are not used for the approximation. This is due to the fact that in practice, the first samples are often not representative for the stationary distribution since they usually depend strongly on the chosen initial value $\boldsymbol{\theta}^{(0)}$ of the chain. The set $\{1, \ldots, r_0\}$ is called *burn-in period*.

There are two classical approaches to construct a Markov chain with stationary distribution $p(\boldsymbol{\theta}|\boldsymbol{x})$. The first one is the Gibbs sampler which was introduced by Geman and Geman (1984) and made popular for statistics by the works of Gelfand and Smith (1990). Assume we have a d dimensional parameter vector $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_d)'$ and data $\boldsymbol{x}$. The Gibbs sampler generates a sample from the posterior distribution $p(\boldsymbol{\theta}|\boldsymbol{x})$ by sequentially generating new values for each component θ_j of $\boldsymbol{\theta}$, using all other components $\boldsymbol{\theta}_{-j} := (\theta_1, \ldots, \theta_{j-1}, \theta_{j+1}, \ldots, \theta_d)'$ and the data $\boldsymbol{x}$. More precisely, this algorithm uses the *full conditional densities*

$$p(\theta_j|\boldsymbol{\theta}_{-j}, \boldsymbol{x}) \quad (j = 1, \ldots, d)$$

to generate samples for each component of $\boldsymbol{\theta}$. The whole procedure is provided in Algorithm 2.1. The full conditional distributions required by the Gibbs sampler may be of similar complexity as the posterior itself. In those cases, it might be a hard task to get samples from them, and thus the Gibbs sampler brings hardly any advantage. Here, one can use the Metropolis-Hastings algorithm developed by Hastings (1970) in extension of the work of Metropolis et al. (1953). This algorithm requires only knowledge of the posterior density up to a proportional constant, as provided in relation (2.4).

Algorithm 2.1 Gibbs sampler

INPUT initial values $\boldsymbol{\theta}^{(0)}$, data $\boldsymbol{x}$, number of MCMC iterations m
OUTPUT sample $\boldsymbol{\theta}^{(1)}, \ldots, \boldsymbol{\theta}^{(m)}$ of the posterior distribution

FOR $r := 1, \ldots, m$ **DO**
  FOR $j := 1, \ldots, d$ **DO**
    Draw a sample $\theta_1^{(r)}$ from $p(\theta_1 | \theta_2^{(r-1)}, \ldots, \theta_d^{(r-1)}, \boldsymbol{x})$
    Draw a sample $\theta_2^{(r)}$ from $p(\theta_2 | \theta_1^{(r)}, \theta_3^{(r-1)}, \ldots, \theta_d^{(r-1)}, \boldsymbol{x})$
    $\vdots$
    Draw a sample $\theta_d^{(r)}$ from $p(\theta_d | \theta_1^{(r)}, \ldots, \theta_{d-1}^{(r)}, \boldsymbol{x})$
  END FOR
END FOR

The Metropolis-Hastings algorithm works as follows: In step r, a sample $\boldsymbol{\theta}^{prop}$ is drawn from a proposal distribution $q(\cdot|\boldsymbol{\theta}^{(r-1)})$ which depends on the current chain value $\boldsymbol{\theta}^{(r-1)}$. The new value is taken with an acceptance probability depending on the posterior density and the proposal density, defined as

$$p_{acc} = \min\left(1, \frac{p(\boldsymbol{\theta}^{prop}|\boldsymbol{x}) q(\boldsymbol{\theta}^{(r-1)}|\boldsymbol{\theta}^{prop})}{p(\boldsymbol{\theta}^{(r-1)}|\boldsymbol{x}) q(\boldsymbol{\theta}^{prop}|\boldsymbol{\theta}^{(r-1)})}\right) \tag{2.5}$$

On the other hand, the sampled value is rejected with probability $1 - p_{acc}$, such that the chain stays at the current value. The Metropolis-Hastings sampler is summarized in Algorithm 2.2.

Algorithm 2.2 Metropolis Hastings algorithm

INPUT initial values $\boldsymbol{\theta}^{(0)}$, data $\boldsymbol{x}$, number of MCMC iterations m
OUTPUT sample $\boldsymbol{\theta}^{(1)}, \ldots, \boldsymbol{\theta}^{(m)}$ of the posterior distribution

FOR $r := 1, \ldots, m$ **DO**
  Draw a sample $\boldsymbol{\theta}^{prop}$ from $q(\cdot|\boldsymbol{\theta}^{(r-1)})$
  Calculate acceptance probability p_{acc} from (2.5)
  Draw a sample x from $Uniform(0, 1)$
  IF $p_{acc} \leq x$ **THEN**
    $\boldsymbol{\theta}^{(r)} := \boldsymbol{\theta}^{(prop)}$
  ELSE
    $\boldsymbol{\theta}^{(r)} := \boldsymbol{\theta}^{(r-1)}$
  END IF
END FOR

The definition of the acceptance probability guarantees the convergence of the created chain towards the posterior distribution $p(\boldsymbol{\theta}|\boldsymbol{x})$. Thus, a nice feature of the Metropolis-Hastings algorithm is that the convergence holds regardless of the choice of the proposal

distribution. However, a bad choice of the proposal may lead to a high number of rejections and therefore to slow convergence.

As alternative to updating all components of $\boldsymbol{\theta}$ at once, one can also divide $\boldsymbol{\theta}$ into blocks and/or single components and apply sequentially Metropolis Hastings steps on each block/component of $\boldsymbol{\theta}$. Furthermore, it is possible to apply the Gibbs sampler to some selected components of $\boldsymbol{\theta}$ and the Metropolis Hastings steps to the remaining ones. This can be helpful in a situation when the full conditional densities are available only for a part of the parameter vector $\boldsymbol{\theta}$. A chain created by an MCMC algorithm which uses both Gibbs and Metropolis Hastings sampling is called a *hybrid chain*. We will use the latter method for our MCMC algorithm.

2.6 Copula and measures of Dependence

2.6.1 Modeling dependence with copulas

Copulas are functions that describe the connection between the joint distribution of random variables and their marginal distributions. The *copula* concept was developed by Abe Sklar in his revolutionary work *Fonctions de répartition à n dimensions et leurs marges*. This concept allows us on the one hand to describe dependence of random variables independently from their marginal distributions and on the other hand to construct joint distributions by separately defining dependence and marginal behavior. In our introduction of copulas, we follow Nelsen (1999).

Definition A *(bivariate) copula* is a function C from $[0,1]^2$ to $[0,1]$ with the following properties:

(i) For every $u, v \in [0,1]$,
$$C(u,0) = 0 = C(0,v)$$

(ii) For every $u, v \in [0,1]$,
$$C(u,1) = u \text{ and } C(1,v) = v$$

(iii) For every $u_1, u_2, v_1, v_2 \in [0,1]$ with $u_1 \leq u_2$ and $v_1 \leq v_2$,
$$C(u_2,v_2) - C(u_2,v_1) - C(u_1,v_2) + C(u_1,v_1) \geq 0$$

We concentrate on the bivariate case, i. e. with copulas we always mean bivariate copulas fulfilling the properties above. Of course, a definition for an arbitrary dimension d can also be easily derived, but we will not need that in this thesis.

At next, we will introduce the important *Sklar's theorem*, which first appeared in Sklar (1959). It allows us to divide a bivariate distribution function into a copula and its marginal distribution functions as well as to construct a bivariate distribution function from a copula and two univariate distribution functions.

Theorem *(Sklar's theorem)* Let H be a joint distribution function with margins F and G. Then there exists a copula C such that for all $x, y \in \overline{\mathbb{R}} := \mathbb{R} \cup \{-\infty, +\infty\}$:

$$H(x, y) = C(F(x), G(y)) \tag{2.6}$$

If G and F are continuous, then C is unique; otherwise, C is uniquely determined on the product space of the range of F and the range of G. Conversely, if C is a copula and F and G are distribution functions, then the function H defined by (2.6) is a joint distribution function with margins F and G.

Proof. See Nelsen (1999), p. 18 □

It is possible to derive a similar result for density functions instead of distribution functions. For that purpose we will look at the partial derivatives of a copula, which we will use to define a copula density and to see when such a density exists. The word "almost" is used in the sense of the Lebesgue measure.

Theorem Let C be a copula. For any $v \in [0, 1]$, the partial derivative $\frac{\partial C(u,v)}{\partial u}$ exists for almost all u, and for such v and u it holds

$$0 \leq \frac{\partial}{\partial u} C(u, v) \leq 1$$

Similarly, for any $u \in [0, 1]$, the partial derivative $\frac{\partial C(u,v)}{\partial v}$ exists for almost all v, and for such u and v it holds

$$0 \leq \frac{\partial}{\partial v} C(u, v) \leq 1$$

Furthermore, the functions $u \mapsto \frac{\partial C(u,v)}{\partial v}$ and $v \mapsto \frac{\partial C(u,v)}{\partial u}$ are defined and nondecreasing almost everywhere on $[0, 1]$.

Proof. See Nelsen (1999), p. 11 □

Theorem Let C be a copula. If $\frac{\partial C(u,v)}{\partial v}$ and $\frac{\partial^2 C(u,v)}{\partial u \partial v}$ are continuous on $[0, 1]^2$ and $\frac{\partial C(u,v)}{\partial u}$ exists for all $u \in (0, 1)$ when $v = 0$, then $\frac{\partial C(u,v)}{\partial u}$ and $\frac{\partial^2 C(u,v)}{\partial v \partial u}$ exist in $(0, 1)^2$ and $\frac{\partial^2 C(u,v)}{\partial v \partial u} = \frac{\partial^2 C(u,v)}{\partial u \partial v}$.

Proof. See Seeley (1961) □

For bivariate distribution functions $H(x, y)$, the density is equal to the twice partial differentiation $\frac{\partial^2 H(x,y)}{\partial x \partial y}$. We will use the same relationship for defining the density of a copula C.

Definition Let C be a twice partial differentiable copula.
The function $c : [0, 1]^2 \mapsto [0, 1]$ defined by

$$c(u, v) = \frac{\partial^2 C(u, v)}{\partial u \partial v}$$

is called the density of the copula C.

With Sklar's theorem and by using the same notation as above, we can use arbitrary marginal densities $f(x)$ and $g(y)$ to construct a joint density $h(x, y)$ from the copula density $c(F(x), G(y))$ and vice versa:

$$\begin{aligned} h(x,y) &= \frac{\partial^2 H(x,y)}{\partial x \partial y} \overset{(2.6)}{=} \frac{\partial^2 C(F(x),G(y))}{\partial x \partial y} = \frac{\partial^2 C(F(x),G(y))}{\partial F(x) \partial G(y)} \frac{\partial F(x)}{\partial x} \frac{\partial G(y)}{\partial y} \\ &= c(F(x),G(y)) f(x) g(y) \end{aligned} \tag{2.7}$$

which means that a copula density is constructed by

$$c(u,v) = \frac{h\left(F^{-1}(u), G^{-1}(v)\right)}{f\left(F^{-1}(u)\right) g\left(F^{-1}(v)\right)}.$$

2.6.2 Measures of dependence

Before we introduce the Gauss copula, we discuss different types of correlations. Reference for this part is Kurowicka and Cooke (2006).

Definition *Product moment correlation.* The product moment correlation of random variables X, Y with finite expectations $E(X), E(Y)$ and finite variances σ_X^2, σ_Y^2, is

$$\rho(X,Y) = \frac{E(XY) - E(X)E(Y)}{\sigma_X \sigma_Y}$$

In this thesis, we mean by the single word "correlation" always the product moment correlation. For random variables $X_1, \ldots, X_d$ and $j, k \in \{1, \ldots, d\}$ we write $\rho_{jk} := \rho(X_j, X_k)$. The matrix $(\rho_{ij})_{j,k=1,\ldots,d}$ is called correlation matrix and denoted by R in most cases.

Definition *Partial correlation.* Let $X_1, \ldots, X_d$ be random variables with zero mean and standard deviations $\sigma_1 = \ldots = \sigma_d = 1$. Let the numbers $b_{12;3,\ldots,d}, \ldots, b_{1d;2,\ldots,d-1}$ minimize

$$E\left((X_1 - b_{12;3,\ldots,d} X_2 - \ldots - b_{1d;2,\ldots,d-1} X_d)^2\right)$$

Then the partial correlation $\rho_{12;3,\ldots,d}$ is defined as

$$\rho_{12;3,\ldots,d} = \operatorname{sgn}\left(b_{12;3,\ldots,d}\right) \sqrt{b_{12;3,\ldots,d} b_{21;3,\ldots,d}}$$

We will mainly use an equivalent definition:

$$\rho_{12;3,\ldots,d} = \frac{\det\left(R_{-(1,2)}\right)}{\sqrt{\det\left(R_{-(1,1)}\right) \det(R_{-(2,2)})}} \tag{2.8}$$

where $R_{-(j,k)}$ denotes the correlation matrix R reduced by the jth row and the kth column.

The partial correlation $\rho_{12;3,\dots,d}$ can be interpreted as the correlation between the orthogonal projections of X_1 and X_2 on the plane orthogonal to the space spanned by $X_3, \dots, X_d$. Yule and Kendall (1965) proposed a recursive formula to calculate partial correlations from the correlation matrix R:

$$\rho_{12;3,\dots,d} = \frac{\rho_{12;3,\dots,d-1} - \rho_{1d;3,\dots,d-1}\rho_{2d;3,\dots,d-1}}{\sqrt{1-\rho^2_{1d;3,\dots,d-1}}\sqrt{1-\rho^2_{2d;3,\dots,d-1}}} \tag{2.9}$$

We will need both formulas later when we derive the correlation matrix R out of a certain set of partial correlations. The last type of correlation that we consider is the conditional correlation.

Definition *Conditional correlation.* Let $X_1, \dots, X_d$ be random variables with finite expectations $E(X_1), \dots E(X_d)$ and finite variances $\sigma^2_{X_1}, \dots \sigma^2_{X_d}$. The conditional correlation of X_1, X_2 given $X_3, \dots, X_d$

$$\begin{aligned}\rho_{12|3,\dots,d} &= \rho(X_1|X_3, \dots, X_d, X_2|X_3, \dots, X_d) \\ &= \frac{E\left(X_1X_2|X_3, \dots, X_d\right)}{\sigma(X_1|X_3, \dots, X_d)\,\sigma(X_2|X_3, \dots, X_d)}\end{aligned}$$

is the (product moment) correlation computed with the conditional distribution of X_1 and X_2 given $X_3, \dots, X_d$.

By exchanging indices, one can derive arbitrary partial correlations $\rho_{jk;I}$ and conditional correlations $\rho_{jk|I}$ for $j, k \in \{1, \dots, d\}$ and $I \subseteq \{1, \dots, d\} \setminus \{j, k\}$ from the definitions above. For consistency, we will permanently use the notation $\rho_{jk;\emptyset} := \rho_{jk|\emptyset} := \rho_{jk}$.

Theorem For any d-dimensional normal distribution, the partial correlation $\rho_{jk;I}$ is equal to the conditional correlation $\rho_{jk|I}$ for all $j, k \in \{1, \dots, d\}$ and $I \in \{1, \dots, d\} \setminus \{j, k\}$.

Proof. See Baba, Shibata, and Sibuya (2004) □

Since our model will be based on the multivariate normal distribution, we will later often use partial and conditional correlations as synonyms.

2.6.3 Gauss copula and the bivariate normal distribution

Definition *(Bivariate) Gauss copula.* Let $\Phi_{2,\rho}(\cdot,\cdot)$ denote the cumulative distribution function of the bivariate standard normal distribution with correlation ρ, let $\Phi(\cdot)$ denote the cdf of the univariate standard normal distribution and $\Phi^{-1}(\cdot)$ its inverse. With the use of Sklar's theorem, we define the Gauss (or normal) copula with parameter $\rho \in [-1, 1]$ as

$$C(u, v|\rho) = \Phi_{2,\rho}\left(\Phi^{-1}(u), \Phi^{-1}(v)\right) \tag{2.10}$$

The density of a Gauss copula with parameter ρ is defined as

$$c(u, v|\rho) = \frac{1}{\sqrt{1-\rho^2}} \exp\left\{-\frac{\rho^2(u^2+v^2) - 2\rho uv}{2(1-\rho^2)}\right\} \tag{2.11}$$

over the support $[0,1]^2$.

Since the Gauss copula is derived from the normal distribution, it is no surprise that the density of an arbitrary bivariate normal distribution can be constructed by the product of a Gauss copula density and two univariate normal densities and vice versa. This is shown by the following theorem.

Theorem Let $c(\cdot,\cdot|\rho)$ denote the density of a Gauss copula, $\Phi_{\mu,\sigma^2}(\cdot)$ the cdf of a univariate normal distribution with mean μ and variance σ^2 and $\varphi_{\mu,\sigma^2}(\cdot)$ its density function.
Let $\boldsymbol{\mu} = (\mu_1, \mu_2)' \in \mathbb{R}^2$, $\boldsymbol{\sigma}^2 = (\sigma_1^2, \sigma_2^2)' \in (0,\infty)^2$ and $\rho \in (-1,1)$.

Then the following statements are equivalent:

- $(X_1, X_2)' \sim \mathcal{N}_2(\boldsymbol{\mu}, \Sigma)$ with $\Sigma := \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}$
- $X_1 \sim \mathcal{N}(\mu_1, \sigma_1^2)$, $X_2 \sim \mathcal{N}(\mu_2, \sigma_2^2)$ with joint density

$$f(x_1, x_2) = c\left(\Phi_{\mu_1,\sigma_1^2}(x_1), \Phi_{\mu_2,\sigma_2^2}(x_2)\middle|\rho\right) \varphi_{\mu_1,\sigma_1^2}(x_1)\varphi_{\mu_2,\sigma_2^2}(x_2) \tag{2.12}$$

Proof. We use the notation $\Phi(\cdot)$, $\Phi^{-1}(\cdot)$, $\varphi(\cdot)$ for the cumulative distribution function, its inverse and the density function of a univariate standard normal distribution and we write $\Phi_{2,\rho}(\cdot,\cdot)$, $\varphi_{2,\rho}(\cdot,\cdot)$ for the cumulative distribution function and density function of a bivariate standard normal distribution with correlation ρ.

We see that

$$\begin{aligned}
&c\left(\Phi_{\mu_1,\sigma_1^2}(x_1), \Phi_{\mu_2,\sigma_2^2}(x_2)\middle|\rho\right) \varphi_{\mu_1,\sigma_1^2}(x_1)\varphi_{\mu_2,\sigma_2^2}(x_2) \\
&= \frac{\partial C}{\partial \Phi_{\mu_1,\sigma_1^2} \partial \Phi_{\mu_2,\sigma_2^2}} \left(\Phi_{\mu_1,\sigma_1^2}(x_1), \Phi_{\mu_2,\sigma_2^2}(x_2)\middle|\rho\right) \varphi_{\mu_1,\sigma_1^2}(x_1)\varphi_{\mu_2,\sigma_2^2}(x_2) \\
&= \frac{\partial C}{\partial x_1 \partial x_2} \left(\Phi_{\mu_1,\sigma_1^2}(x_1), \Phi_{\mu_2,\sigma_2^2}(x_2)\middle|\rho\right)
\end{aligned}$$

where we applied in the last step the chain rule for differentiation.

Since $\Phi_{\mu,\sigma^2}(x) = \Phi(\frac{x-\mu}{\sigma})$, the expression above is equal to

$$
\begin{aligned}
&\frac{\partial C}{\partial x_1 \partial x_2}\left(\Phi\left(\frac{x_1-\mu_1}{\sigma_1}\right), \Phi\left(\frac{x_2-\mu_2}{\sigma_2}\right)\Big|\rho\right) \\
&\overset{(2.10)}{=} \frac{\partial}{\partial x_1 \partial x_2}\Phi_{2,\rho}\left(\Phi^{-1}\left(\Phi\left(\frac{x_1-\mu_1}{\sigma_1}\right)\right), \Phi^{-1}\left(\Phi\left(\frac{x_2-\mu_2}{\sigma_2}\right)\right)\right) \\
&= \frac{\partial}{\partial x_1 \partial x_2}\Phi_{2,\rho}\left(\frac{x_1-\mu_1}{\sigma_1}, \frac{x_2-\mu_2}{\sigma_2}\right) \\
&= \varphi_{2,\rho}\left(\frac{x_1-\mu_1}{\sigma_1}, \frac{x_2-\mu_2}{\sigma_2}\right)\frac{1}{\sigma_1}\frac{1}{\sigma_2} \\
&= \frac{1}{2\pi\sqrt{1-\rho^2}}\exp\left\{\frac{1}{2(1-\rho^2)}\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2 + \left(\frac{x_2-\mu_2}{\sigma_2}\right)^2 - 2\rho\left(\left(\frac{x_1-\mu_1}{\sigma_1}\right)\left(\frac{x_2-\mu_2}{\sigma_2}\right)\right)\right\} \\
&\quad \cdot \frac{1}{\sigma_1}\frac{1}{\sigma_2}
\end{aligned}
$$

which is equal to the density of $\mathcal{N}_2(\boldsymbol{\mu}, \boldsymbol{\sigma}^2)$ given in (2.1) □

2.7 Pair-copula constructions and vines

2.7.1 Modeling dependence with bivariate copulas

Bivariate copulas can be used to express or construct a multivariate distribution by specifying the dependence and conditional dependence of selected pairs of random variables and all marginal distribution functions. The idea was developed by Bedford and Cooke (2002) based on the work of Joe (1996). Our presentation of the concept follows Aas et al. (2009).

Consider a vector $\boldsymbol{X} = (X_1, \ldots, X_d)'$ of random variables with a joint density function $f(x_1, \ldots, x_d)$. Using conditional densities, we can factorize the joint density as

$$f(x_1, \ldots, x_d) = f_d(x_d) \cdot f(x_{d-1}|x_d) \cdot f(x_{d-2}|x_{d-1}, x_d) \cdots f(x_1|x_2, \ldots, x_d) \tag{2.13}$$

and this decomposition is unique up to a re-labeling of the variables. We now want to use this decomposition to express the joint density as product of copula densities and the marginal densities. For the bivariate case, we already know from (2.7) that

$$f(x_1, x_2) = c_{12}(F_1(x_1), F_2(x_2))f_1(x_1)f_2(x_2) \tag{2.14}$$

where c_{12} is the appropriate *pair-copula density* for the pair of transformed variables $F_1(x_1)$ and $F_2(x_2)$. For a conditional density $f(x_1|x_2)$, it follows that

$$f(x_1|x_2) = \frac{f(x_1, x_2)}{f_2(x_2)} = c_{12}(F_1(x_1), F_2(x_2))f_1(x_1) \tag{2.15}$$

We can also apply the representation (2.15) when we are conditioning on further variables.

Thus, for three random variables X_1, X_2 and X_3 we see that

$$f(x_1|x_2,x_3) = \frac{f(x_1,x_2|x_3)}{f(x_2|x_3)} = c_{12|3}(F(x_1|x_3),F(x_2|x_3))f(x_1|x_3)$$

for the appropriate pair-copula density $c_{12|3}$, applied to the transformed variables $F(x_1|x_3)$ and $F(x_2|x_3)$. We can further decompose $f(x_1|x_3)$ in the way of (2.15), which leads to

$$f(x_1|x_2,x_3) = c_{12|3}(F(x_1|x_3),F(x_2|x_3))c_{13}(F_1(x_1),F_3(x_3))f_1(x_1)$$

So with the decomposition (2.13), we can factorize the joint density of X_1, X_2 and X_3 in the form

$$\begin{aligned} f(x_1,x_2,x_3) &= f_3(x_3)f(x_2|x_3)f(x_1|x_2,x_3) \\ &= c_{12|3}(F(x_1|x_3),F(x_2|x_3))c_{13}(F_1(x_1),F_3(x_3))c_{23}(F_2(x_2),F_3(x_3)) \\ &\quad \cdot f_1(x_1)f_2(x_2)f_3(x_3) \end{aligned}$$

where we used $f(x_2|x_3) = c_{23}(F_2(x_2),F_3(x_3))f_2(x_2)$. One can use similar decompositions to get other pair-copula constructions like

$$\begin{aligned} f(x_1,x_2,x_3) &= c_{13|2}(F(x_1|x_2),F(x_3|x_2))c_{12}(F_1(x_1),F_2(x_2))c_{23}(F_2(x_2),F_3(x_3)) \\ &\quad \cdot f_1(x_1)f_2(x_2)f_3(x_3) \end{aligned}$$

where the copula density $c_{13|2}$ is different from the density $c_{12|3}$ used in the previous construction.

For the general case of random variables $X_1,\dots,X_d$, one can obviously obtain many different pair-copula constructions for the joint density $f(x_1,\dots,x_d)$ by using the factorization (2.13) and applying the formula

$$f(x|\boldsymbol{v}) = c_{xv_j}(F(x|\boldsymbol{v}_{-j}),F(v_j|\boldsymbol{v}_{-j}))f(x|\boldsymbol{v}_{-j})$$

where $\boldsymbol{v}$ is a vector of dimension $k < d$, $j \in \{1,\dots,k\}$ and $\boldsymbol{v}_{-j}$ denotes the vector $\boldsymbol{v}$ reduced by the component v_j. To calculate the marginal conditional distribution functions $F(x|\boldsymbol{v})$ involved in the pair-copula construction, one uses the relationship

$$F(x|\boldsymbol{v}) = \frac{\partial C_{xv_j|\boldsymbol{v}_{-j}}\left(F(x|\boldsymbol{v}_{-j}),F(v_j|\boldsymbol{v}_{-j})\right)}{\partial F(v_j|\boldsymbol{v}_{-j})}$$

proved by Joe (1996), where $C_{xv_j|\boldsymbol{v}_{-j}}$ is a copula. So for univariate v, we get

$$F(x|v) = \frac{\partial C_{xv}\left(F(x),F(v)\right)}{\partial F(v)} =: h(F(x),F(v),\boldsymbol{\theta}_{xv}) \tag{2.16}$$

where $\boldsymbol{\theta}_{xv}$ are the parameters of the copula C_{xv}. We use the notation of the h-function later when we calculate the likelihood of a certain pair-copula construction.

The h-function depends on the type of the associated copula. For the Gauss copula, the h-function can be easily derived: Recall that the Gauss copula is defined by

$$C(u_1, u_2|\rho) = \Phi_{2,\rho}\left(\Phi^{-1}(u_1), \Phi^{-1}(u_2)\right)$$

For random variables $(X, V)' \sim \mathcal{N}_2\left(\mathbf{0}, \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}\right)$, the conditional distribution of X given $V = v$ is $\mathcal{N}(\rho v, 1-\rho^2)$, so the h-function is given by

$$h(u_1, u_2, \rho) = \Phi\left(\frac{\Phi^{-1}(u_1) - \rho\Phi^{-1}(u_2)}{\sqrt{1-\rho^2}}\right)$$

2.7.2 Vines

The number of possible pair-copula constructions for a multivariate distribution increases significantly with the dimension. For example, there are 240 possible decompositions for a 5 dimensional distribution. Bedford and Cooke (2002) have developed a structure that helps to organize them, the regular vine. In this thesis, we concentrate on an interesting special case, the *D-vine*, which was introduced by Kurowicka and Cooke (2004). Essentially, we follow again the presentation in Aas et al. (2009).

A *vine* $\mathcal{V}$ on d variables consists of $d-1$ nested trees $T_1, \ldots, T_{d-1}$ that are specified as follows: For each $j \in \{1, \ldots, d-1\}$, tree T_j consists of $d-j+1$ nodes (and thus $d-j$ edges), and each edge of tree T_j becomes a node in tree T_{j+1}. A *regular vine* $\mathcal{V}$ is a vine where two nodes in tree T_{j+1} may be joined by an edge only if they share a common node in tree T_j. The complete set of edges of $\mathcal{V}$ is called constraint set and denoted by $\mathcal{CV}$. Each of the $\frac{n(n-1)}{2}$ edges of a regular vine corresponds to a bivariate copula density in a pair-copula construction. Adding the marginal densities, we get the complete decomposition.

A *D-vine* is a regular vine where each node is connected to at most 2 other nodes. Figure 2.1 shows as example a D-vine on 4 variables, consisting of 3 trees and 6 edges. We can denote the edges of each vine by $(jk|j+1, \ldots, k-1)$, where $1 \leq j < k \leq d$. The edge $(jk|j+1, \ldots, k-1) \in \mathcal{CV}$ corresponds to a copula density $c_{jk|j+1,\ldots,k-1}$ in a pair-copula construction.

The corresponding pair-copula density of the D-vine may be written as

$$\begin{aligned} &f(x_1, \ldots, x_d) \\ &\quad = \prod_{k=1}^{d} f_k(x_k) \prod_{l=1}^{d-1} \prod_{j=1}^{d-l} c_{j,j+l|j+1,\ldots,j+l-1}\left(F(x_j|x_{j+1}, \ldots, x_{j+l-1}), F(x_{j+l}|x_{j+1}, \ldots, x_{j+l-1})\right) \end{aligned} \tag{2.17}$$

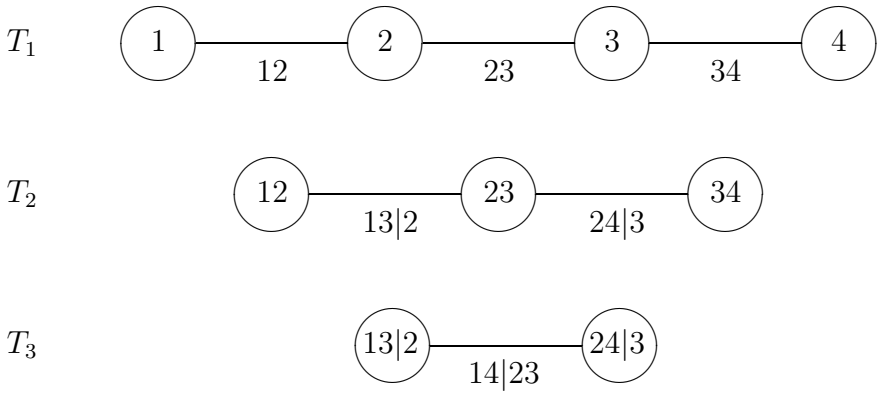

Figure 2.1: Illustration of a D-vine on 4 variables, consisting of trees T_1, T_2, T_3. Each of the 6 edges $(jk|j+1,\dots,k-1)$ corresponds to a bivariate copula density

To calculate the conditional distribution functions occurring in the likelihood (2.17) in an efficient way, we iteratively use the relationships denoted in equation (2.16) and the corresponding h-functions. To see how this works, we look at three random variables X_1, X_2, X_3 and calculate the conditional distribution function $f(x_1|x_2, x_3)$. To simplify the demonstration, we omit the copula parameters in the h-functions.

$$\begin{aligned} f(x_1|x_2,x_3) &\overset{(2.16)}{=} \frac{\partial C_{13|2}(F(x_1|x_2), F(x_3|x_2))}{\partial F(x_3|x_2)} \\ &= h(F(x_1|x_2), F(x_3|x_2)) \\ &= h\left(\frac{\partial C_{12}(F_1(x_1), F_2(x_2))}{\partial F_2(x_2)}, \frac{\partial C_{23}(F_2(x_2), F_3(x_3))}{\partial F_2(x_2)}\right) \\ &= h(h(F_1(x_1), F_2(x_2)), h(F_3(x_3), F_2(x_2))) \\ &= h(h(u_1, u_2), h(u_3, u_2)) \end{aligned}$$

where $u_j := F_j(x_j)$ for $j = 1, 2, 3$. In the same way, we get all other values of conditional distribution functions in the D-vine likelihood, using only the h-functions, the copula parameters and $u_1 := F_1(x_1), \dots, u_d := F_d(x_d)$. Algorithm 2.3 provides the calculation of the log-likelihood for uniform marginal distributions. To simplify the description, we use the notation

$$L(x, v|\boldsymbol{\theta}) := log(c(x, v|\boldsymbol{\theta}))$$

where $\boldsymbol{\theta}$ is the parameter of the copula density c. Since $u_1, \dots, u_d$ follow a $Uniform(0, 1)$ distribution, we can calculate the copula part of the D-vine log-likelihood with Algorithm 2.3. The whole expression is then determined by adding the logarithms of the marginal densities.

Algorithm 2.3 Log-likelihood evaluation for a D-vine with uniformly distributed marginals

INPUT Data $u_1, \dots, u_d \in (0,1)$, copula parameters $\{\boldsymbol{\theta}_{jk|j+1,\dots,k-1}, 1 \le j < k \le d\}$
OUTPUT Log-likelihood L_{vine}

L_{vine}:=0
FOR $j := 1, \dots, d-1$ **DO**
  $v_{0,j} := u_j$
END FOR
FOR $j := 1, \dots, d-1$ **DO**
  $L_{vine} := L_{vine} + L(v_{0,j}, v_{0,j+1}|\boldsymbol{\theta}_{j,j+1})$
END FOR
$v_{1,1} := h(v_{0,1}, v_{0,2}, \theta_{12})$
FOR $j := 1, \dots, d-3$ **DO**
  $v_{1,2j} := h(v_{0,j+2}, v_{0,j+1}, \boldsymbol{\theta}_{j+1,j+2})$
  $v_{1,2j+1} := h(v_{0,j+1}, v_{0,j+2}, \boldsymbol{\theta}_{j+1,j+2})$
END FOR
$v_{1,2d-4} := h(v_{0,d}, v_{0,d-1}, \boldsymbol{\theta}_{d-1,d})$
FOR $j := 2, \dots, d-1$ **DO**
  FOR $k := 1 \dots, d-1$ **DO**
    $L_{vine} := L_{vine} + L(v_{j-1,2k-1}, v_{j-1,2k}|\boldsymbol{\theta}_{k,k+j|k+1,\dots,k+j-1})$
  END FOR
  IF $j = d-1$ **THEN**
    STOP
  END IF
  $v_{j,1} := h(v_{j-1,1}, v_{j-1,2}, \boldsymbol{\theta}_{1,1+j|2,\dots,j})$
  IF $d > 4$ **THEN**
    FOR $k := 1, \dots, d-j-2$ **DO**
      $v_{j,2k} := h(v_{j-1,2k+2}, v_{j-1,2k+1}, \boldsymbol{\theta}_{k+1,k+j+1|k+2,\dots,k+j})$
      $v_{j,2k+1} := h(v_{j-1,2k+1}, v_{j-1,2k+2}, \boldsymbol{\theta}_{k+1,k+j+1|k+2,\dots,k+j})$
    END FOR
  END IF
  $v_{j,2d-2j-2} := h(v_{j-1,2d-2j}, v_{j-1,2d-2j-1}, \boldsymbol{\theta}_{d-j,d|d-j+1,\dots,d-1})$
END FOR

2.7.3 D-vines and the multivariate normal distribution

In Section 2.6.3 we showed how a bivariate normal distribution can be expressed as product of a Gauss pair-copula density and two univariate normal densities. We see that this corresponds to a pair-copula construction organized on a two dimensional D-vine: The constraint set of this vine is $\mathcal{CV} = \{(1,2)\}$, which leads to the decomposition

$$f(x_1, x_2) = c_{12}\left(F_1(x_1), F_2(x_2)|\boldsymbol{\theta}_{12}\right) f(x_1) f(x_2)$$

If $c_{12}(\cdot,\cdot|\boldsymbol{\theta}_{12})$ is the density of a Gauss copula with parameter $\boldsymbol{\theta}_{12} = \rho_{12}$ and $f(x_1), f(x_2)$ are univariate normal densities, we get our previously considered product (2.12), so we know that the joint density is normal with correlation ρ_{12}.

So the question arises if this relation also holds for D-vines with dimension $d > 2$ and if the copula parameters $\rho_{jk|I}$ with $(jk|I) \in \mathcal{CV}$ can be interpreted as conditional correlations as the notation suggests.
We will see that this is the case.

In the following, we use the notation $j:k$, which we define as $j:k := j, j+1, \ldots, k$ for $j < k$, $j:k := j$ for $j = k$ and $j:k := \emptyset$ for $j > k$.

Theorem Let $\mathcal{V}$ be a d-dimensional D-vine and for each element

$$(jk|j+1:k-1), \quad 1 \le j < k \le d$$

of the D-vine let $c_{jk|j+1:k-1}(\cdot,\cdot|\rho_{jk|j+1:k-1})$ denote the density of a Gauss copula with parameter $\rho_{jk|j+1:k-1} \in (-1,1)$. Let $\Phi(\cdot)$ denote the cdf of the standard normal distribution and $\varphi(\cdot)$ its density function.

Then the following statements are equivalent:

- $(X_1, \ldots, X_d)' \sim \mathcal{N}(\mathbf{0}, R)$ with conditional correlations $\rho_{jk|j+1:k-1}$
- $X_1 \sim \mathcal{N}(0,1), \ldots, X_d \sim \mathcal{N}(0,1)$ with joint density

$$f(x_1, \ldots, x_d) = \prod_{k=1}^{d} \varphi(x_k) \prod_{l=1}^{d-1} \prod_{j=1}^{d-l} c_{j,j+l|j+1:j+l-1}\left(F(x_j|\boldsymbol{x}_{j+1:j+l-1}), F(x_{j+l}|\boldsymbol{x}_{j+1:j+l-1})|\rho_{j,j+l|j+1:j+l-1}\right)$$

Proof. Czado and Min (2009) show how a multivariate joint density can be decomposed into a D-vine density. In our case, the joint density is multivariate normal and is decomposed – as in the bivariate case – into the marginal normal densities and Gauss copula densities whose correlation parameters are equal to the corresponding conditional correlations. The copula parameters $\rho_{j,j+l|j+1:j+l-1} \in (-1,1)$ are conditional correlations and due to the normal distribution equal to the partial correlations $\rho_{j,j+l;j+1:j+l-1}$, which form

a partial correlation specification on a D-vine. Bedford and Cooke (2002) show that for any d-dimensional regular vine $\mathcal{V}$ there exists a one-to-one correspondence between the set of d-dimensional correlation matrices and the set of partial correlation specifications for the vine $\mathcal{V}$ (see Bedford and Cooke (2002), Corollary 7.5). □

2.7.4 Calculating the correlation matrix from a partial correlation D-vine

We have seen in Section 2.6.2 how a partial correlation may be calculated from the correlation matrix R. But we are also interested in how the correlation matrix R can be derived from partial correlations adapted to a D-vine. We developed an algorithm that performs this calculation.

Before we provide the algorithm, we have to do some preparations. At first, we solve the recursive formula (2.9) of Yule and Kendall (1965) for $\rho_{12;3,\dots,d-1}$, such that we get

$$\begin{aligned}\rho_{12;3,\dots,d-1} &= \rho_{1d;3,\dots,d-1}\rho_{2d;3,\dots,d-1} + \rho_{12;3,\dots,d}\sqrt{1-\rho^2_{1d;3,\dots,d-1}}\sqrt{1-\rho^2_{2d;3,\dots,d-1}} \\ &=: g(\rho_{12;3,\dots,d}|\rho_{1d;3,\dots,d-1}, \rho_{2d;3,\dots,d-1}) \end{aligned} \tag{2.18}$$

At next, we need to calculate partial correlations from reordered submatrices of the correlation matrix. For that purpose, we use the following notation:

For $d \in \mathbb{N}$ and $l \leq d$, let $(j_1, \dots, j_l)$ be an ordered subset of $\{1, \dots, d\}$ and R a matrix of dimension $d \times d$. We write $R[j_1, j_2, \dots, j_l]$ for a matrix M of dimension $l \times l$ that satisfies

$$M_{ik} = R_{j_i, j_k} \quad \forall i, k \in \{1, \dots, l\}$$

This means that the rows and columns of R are rearranged according to the order of $(j_1, \dots, j_l)$. If $l < d$, then $R[j_1, \dots, j_l]$ does not contain all entries of R.

For any correlation matrix $R \in \mathbb{R}^{d\times d}$ and ordered subset $(j_1, \dots, j_l)$ of $\{1, \dots, d\}$, we can use $M := R[j_1, \dots, j_l]$ as correlation matrix in formula (2.8) to calculate the partial correlation

$$\rho_{j_1 j_2; j_3, \dots, j_l} = \frac{\det\left(M_{-(1,2)}\right)}{\sqrt{\det\left(M_{-(1,1)}\right)\det(M_{-(2,2)})}} \tag{2.19}$$

Algorithm 2.4 now shows the calculation of the correlation matrix R from a vector of partial correlations adapted to a D-vine. We describe how this algorithm works: At the beginning, all unconditional correlations are taken from the first tree of the D-vine. So we know all correlations of the type ρ_{jk} with $|j-k| < 2$. In the case $d = 2$ we have finished, so we assume in the following that $d > 2$ holds.

In step l, we know all correlations ρ_{jk} with $|j-k| < l$. We take $\rho_{1,1+l;2:l} =: \nu_1$ from the D-vine and calculate the partial correlation $\rho_{1,1+l;2:l-1}$ as follows: At first, we determine the partial correlation $\rho_{l,1+l;2:l-1} := \nu_2$ with formula (2.19) and $M := R[l, 1+l, 2:l-1]$.

Algorithm 2.4 Calculation of the correlation matrix from a partial correlation D-vine

INPUT dimension d
partial correlations $\{\rho_{jk;j+1:k-1}\}$ adapted to a d-dimensional D-vine
OUTPUT correlation matrix R

Create empty matrix R
$diag(R) := (\underbrace{1, \ldots, 1}_{d \text{ times}})'$
FOR $j := 1, \ldots, d-1$ **DO**
  $R_{j,j+1} := R_{j+1,j} := \rho_{j,j+1}$
END FOR
IF $d > 2$ **THEN**
  FOR $l := 2, \ldots, d-1$ **DO**
    FOR $j := 1, \ldots, d-l$ **DO**
      $\nu_1 := \rho_{j,j+l;j+1:j+l-1}$
      FOR $k := 1, \ldots, l-1$ **DO**
        $M := R[j+l-k, j+l, j+1 : j+l-k-1]$
        Use M as correlation matrix and (2.19) to calculate
          $\nu_2 := \rho_{j+l-k,j+l;j+1:j+l-k-1}$
        $\nu_1 := \rho_{j,j+l;j+1:j+l-k-1} = g\left(\nu_1 | \rho_{j,j+l-k;j+1:j+l-k-1}, \nu_2\right)$ from (2.18)
      END FOR
      $R_{j,j+l} := R_{j+l,j} := \nu_1$
    END FOR
  END FOR
END IF

Although the correlation matrix R is not entirely known yet, the entries of M are determined, since M consists only of correlations ρ_{jk} with $|j-k| \leq |(1+l)-2| < l$ that have been ascertained previous to step l.

Afterwards, we take $\rho_{1l;2:l-1}$ from the D-vine and use (2.18) to calculate

$$\rho_{1,1+l;2:l-1} := g\left(\rho_{1,1+l;2:l} | \rho_{1l;2:l-1}, \rho_{l,1+l;2:l-1}\right) = g(\nu_1 | \rho_{1l;2:l-1}, \nu_2)$$

If $l = 2$, we have $\rho_{13} = \rho_{1,1+l;2:l-1}$ and add it to our correlation matrix. Otherwise, for $l > 2$, we redefine the auxiliary variable ν_1 as $\nu_1 := \rho_{1,1+l;2:l-1}$ and calculate $\rho_{l-1,1+l;2:l-2} =: \nu_2$ with (2.19) and $M := R[l-1, 1+l, 2 : l-2]$. Again, all entries of M are known, since $|j-k| \leq |(1+l)-2| < l$ for all $j, k \in \{l-1, 1+l, 2 : l-2\}$. At next, we get $\rho_{1,1+l;2:l-2}$ with (2.18) by evaluating

$$\rho_{1,1+l;2:l-2} := g\left(\rho_{1,1+l;2:l-1} | \rho_{1,l-1;2:l-2}, \rho_{1+l,l-1;2:l-2}\right) = g\left(\nu_1, \rho_{1,l-1;2:l-2}, \nu_2\right).$$

In the same way we determine the partial correlations $\rho_{1,1+l:2:l-3}$, $\rho_{1,1+l:2:l-4}$, etc. until we reach $\rho_{1,1+l} = \rho_{1,1+l;2:l-(l-1)}$ and store this correlation in the correlation matrix R. Summing up, we have determined the first unconditional correlation of step l by successively reducing the conditioned set of the partial correlation $\rho_{1,1+l;2:l}$.

For each $j \in \{2, \ldots, d-l\}$, we perform similar calculations to get the unconditional correlations $\rho_{j,j+l}$ from the partial correlations $\rho_{j,j+l;j+1:j+l-1}$. Thus, at the end of step l, we know all correlations ρ_{jk} with $|j-k| \leq l$. We go on with step $l+1$ until we reach step $d-1$, such that all correlations ρ_{jk} with $|j-k| \leq d-1$ are known, which means that we have the complete correlation matrix.

2.8 Parameter estimation

2.8.1 Estimation methods

There exist many different methods to get estimates for unknown parameters of a statistical model. We will use an MCMC method to get Bayesian estimates. However, to get initial values for the algorithm, it is useful to calculate classical estimates as the *maximum likelihood estimator*, see for example Bickel and Doksum (2001). For the following definitions, assume again that we have a statistical model $(\Omega, \mathcal{F}, \{P_{\boldsymbol{\theta}} | \boldsymbol{\theta} \in \boldsymbol{\Theta}\})$ with unknown parameter vector $\boldsymbol{\theta} \in \boldsymbol{\Theta}$ and an n-dimensional data set $\boldsymbol{x} = (\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n)$ from independent random vectors $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n$ with likelihood

$$f(\boldsymbol{x}|\boldsymbol{\theta}) = \prod_{i=1}^{n} f_i(\boldsymbol{x}_i|\boldsymbol{\theta})$$

Definition The *maximum likelihood estimator* (MLE) for $\boldsymbol{\theta}$ is defined by

$$\hat{\boldsymbol{\theta}} := \arg\max_{\boldsymbol{\theta} \in \boldsymbol{\Theta}} \{f(\boldsymbol{x}|\boldsymbol{\theta})\}$$

We have seen in Section 2.7 that one can divide a multivariate distribution of dimension d into a product of bivariate copula densities and the marginal densities. If we denote the marginal parameters by $\boldsymbol{\theta}^{(mar)}$ and the copula parameters by $\boldsymbol{\theta}^{(cop)}$, and furthermore their associated parameter spaces by $\boldsymbol{\Theta}^{(mar)}$ and $\boldsymbol{\Theta}^{(cop)}$, the whole parameter vector is given by $\boldsymbol{\theta} = (\boldsymbol{\theta}^{(mar)\prime}, \boldsymbol{\theta}^{(cop)\prime})'$. So we get an estimator for $\boldsymbol{\theta}$ by initially maximizing the marginal likelihood which is independent of $\boldsymbol{\theta}^{(cop)}$, so we get an estimate $\hat{\boldsymbol{\theta}}^{(mar)}$ for $\boldsymbol{\theta}^{(mar)}$, and afterwards maximizing the joint likelihood conditional on $\boldsymbol{\theta}^{(mar)} = \hat{\boldsymbol{\theta}}^{(mar)}$. Details on this method can be found in Joe and Xu (1996).

In the following, we denote the jth marginal likelihood of $\boldsymbol{X}_i = (X_{i1}, \ldots, X_{id})'$ by f_{ij}, for $j = 1, \ldots, d$. We can write the marginal parameter vector as $\boldsymbol{\theta}^{(mar)} = (\boldsymbol{\theta}_1^{(mar)\prime}, \ldots, \boldsymbol{\theta}_d^{(mar)\prime})$, where the marginal likelihood f_{ij} depends only on $\boldsymbol{\theta}_j^{(mar)}$ for all $j \in \{1, \ldots, d\}$.

Definition The *inference for margins* (IFM) estimator is $\hat{\boldsymbol{\theta}} = (\hat{\boldsymbol{\theta}}^{(mar)\prime}, \hat{\boldsymbol{\theta}}^{(cop)\prime})'$ where

$$\hat{\boldsymbol{\theta}}_j^{(mar)} := \arg\max_{\boldsymbol{\theta}_j^{(mar)} \in \boldsymbol{\Theta}_j^{(mar)}} \{f_{ij}(x_{ij}|\boldsymbol{\theta}_j^{(mar)})\} \quad \forall j \in \{1, \ldots, d\}$$

$$\hat{\boldsymbol{\theta}}^{(cop)} := \arg\max_{\boldsymbol{\theta}^{(cop)} \in \boldsymbol{\Theta}^{(cop)}} \{f(\boldsymbol{x}|\boldsymbol{\theta}^{(cop)}, \hat{\boldsymbol{\theta}}^{(mar)})\}$$

We will sometimes use empirical estimates for the copula parameters instead of the conditional maximization of the likelihood. In those cases, we will mention this explicitly in the text.

For Bayesian inference, we include prior information in the estimation of the parameter vector. For all cases where the prior distribution is informative, it is convenient to maximize the posterior distribution instead of the likelihood.

Definition The *posterior mode* estimator is defined by

$$\hat{\boldsymbol{\theta}} := \arg\max_{\boldsymbol{\theta}\in\boldsymbol{\Theta}}\{p(\boldsymbol{\theta}|\boldsymbol{x})\}$$

We will use later an approximation for the posterior density by an MCMC algorithm to calculate the posterior mode.

2.8.2 Quality of estimates

When an algorithm is developed to estimate parameters of a model, one is interested in the goodness of the estimation process. The estimates of a good algorithm should be close to the true parameter value θ. We introduce statistics that contain information about the goodness of an estimator and show how they can be estimated if an i.i.d. sample $\{\hat{\theta}^{(1)}, \ldots, \hat{\theta}^{(r)}, r \in \mathbb{N}\}$ of estimates generated by the algorithm is given. Reference for this is Section 2.6 of Kastenmeier (2008) and Venables and Ripley (2003).

Definition The *bias* of an estimator $\hat{\theta}$ for a parameter θ with true value θ_{true} is defined as

$$b(\hat{\theta}) := E(\hat{\theta}) - \theta_{true}$$

Thus, the bias measures the expected deviation of the estimator from the true parameter value. If $b(\hat{\theta}) = 0$, i. e. the expected value of the estimator is equal to the true value, we call $\hat{\theta}$ *unbiased*.

To estimate the bias, we have to estimate the expectation $E(\hat{\theta})$. With our i.i.d. sample $\{\hat{\theta}^{(1)}, \ldots, \hat{\theta}^{(r)}, r \in \mathbb{N}\}$, we can calculate the sample mean

$$\bar{\theta} = \frac{1}{r}\sum_{k=1}^{r}\hat{\theta}^{(k)}$$

which is an unbiased and consistent estimator for $E(\hat{\theta})$ (see e.g. Georgii (2002), pp. 194 and 202).

With the sample mean, we get an estimator for the bias:

$$\hat{b}(\hat{\theta}) := \bar{\theta} - \theta_{true} = \frac{1}{r}\sum_{k=1}^{r}\hat{\theta}^{(k)} - \theta_{true}$$

It is also important to look at the standard error of the estimated bias, which is equal to the standard error of the sample mean, since θ_{true} is constant. The variance of the sample mean is defined by

$$Var\left(\bar{\theta}\right) = Var\left(\frac{1}{r}\sum_{k=1}^{r}\hat{\theta}^{(k)}\right) = \frac{1}{r^2}\sum_{k=1}^{r}Var(\hat{\theta}^{(k)}) = \frac{1}{r}Var(\hat{\theta}) \tag{2.20}$$

where the penultimate equality of (2.20) results from the independence of the sample $\{\hat{\theta}^{(1)}, \ldots, \hat{\theta}^{(r)}, r \in \mathbb{N}\}$. For the variance $Var(\hat{\theta})$, we have the unbiased and consistent estimator (see again Georgii (2002), pp. 194 and 202)

$$s^2(\hat{\theta}) := \frac{1}{r-1}\sum_{k=1}^{r}\left(\hat{\theta}^{(k)} - \bar{\theta}\right)^2 \tag{2.21}$$

which is called *sample variance*. So using (2.20) and (2.21), we get as estimator for the variance of the sample mean and of the estimated bias

$$s^2(\bar{\theta}) = s_b^2(\hat{\theta}) := \frac{1}{r}s^2(\hat{\theta}) = \frac{1}{r(r-1)}\sum_{k=1}^{r}\left(\hat{\theta}^{(k)} - \bar{\theta}\right)^2 \tag{2.22}$$

and their standard errors can be determined by taking the square root of (2.22).

So far, we measured the absolute deviation of the estimator $\hat{\theta}$ from the true value θ_{true}. For the comparison between estimates of parameters with different true values, it makes more sense to compare the relative deviations. For that purpose, we define the relative bias and provide estimates for its value and its standard error.

Definition The *relative bias* of an estimator $\hat{\theta}$ for a parameter θ with true value $\theta_{true} \neq 0$ is defined as

$$rb(\hat{\theta}) := \frac{E(\hat{\theta}) - \theta_{true}}{\theta_{true}} = \frac{b(\hat{\theta})}{\theta_{true}}$$

From this definition, it is straightforward to get an estimator for the relative bias:

$$\widehat{rb}(\hat{\theta}) := \frac{\hat{b}(\hat{\theta})}{\theta_{true}} = \frac{\bar{\theta} - \theta_{true}}{\theta_{true}}$$

The variance of the estimated relative bias can be transformed into

$$Var\left(\widehat{rb}(\hat{\theta})\right) = Var\left(\frac{\hat{b}(\hat{\theta})}{\theta_{true}}\right) = \frac{1}{\theta_{true}^2} Var\left(\hat{b}(\hat{\theta})\right)$$

So an estimator for the variance of the relative bias is given by

$$s_{rb}^2(\hat{\theta}) := \frac{1}{\theta_{true}^2} s_b^2(\hat{\theta}) \overset{(2.22)}{=} \frac{1}{\theta_{true}^2 r(r-1)} \sum_{k=1}^{r} \left(\hat{\theta}^{(k)} - \bar{\theta}\right)^2 \tag{2.23}$$

and the associated standard error is $s_{rb}(\hat{\theta}) := \sqrt{s_{rb}^2(\hat{\theta})}$.

2.9 The capital asset pricing model (CAPM)

The capital asset pricing model (CAPM) is used to describe the rate of return of an asset like a stock or a bond in an equilibrium market when this asset is added to a well-diversified portfolio. The idea was developed among others in the works of Sharpe (1964) and Lintner (1965) based on the portfolio optimization theory developed by Markowitz (see Markowitz (1952)).

Imagine that an investor can allocate her wealth into $d \in \mathbb{N}$ assets $1, \ldots, d$ with returns $Y_1, \ldots, Y_d$. She wants to choose an optimal portfolio $\boldsymbol{w} = (w_1, \ldots, w_d)'$ with weights $w_j \in \mathbb{R}\ \forall j \in \{1, \ldots, d\}, \sum_{i=1}^{d} w_i = 1$ and return

$$Y_p = \sum_{j=1}^{d} w_j Y_j$$

In the Markowitz framework, an investor takes into account only the expected return $E(Y_p)$ and the variance $Var(Y_p)$ (or equivalently its standard deviation). She accepts to bear a higher risk – represented by the standard deviation – only in exchange for a higher expected return. If a riskless asset exists, that is an asset Y_0 with expected return $E(Y_0) = r_0$ and variance $Var(Y_0) = 0$, and if the investor uses a so-called quadratic utility function, Markowitz shows that the optimal portfolio always consists of a proportion of the riskless asset and of one other portfolio which is called the tangency portfolio.

The capital asset pricing model provides different assumptions on the available assets on the market and on the behavior of the market participants. Applying Markowitz optimization, one sees that the tangency portfolio is equal to a portfolio of all available assets weighted by their market value. This portfolio is therefore also called the market portfolio.

As result of the CAPM assumptions, the expected return of each available asset $Y_1, \ldots, Y_d$ in a market equilibrium is given by

$$E(Y_j) = r_0 + \beta_j(E(Y_M) - r_0), \quad j \in \{1, \ldots, d\} \tag{2.24}$$

where Y_M is the return of the market portfolio. So the expected return of each asset is the riskless rate of return plus a risk premium depending on the value of β_j and the expected return of the market. We can rearrange (2.24) to

$$E(Y_j) - r_0 = \beta_j E(Y_M - r_0), \quad j \in \{1, \ldots, d\}$$

meaning that the expected excess return of each asset Y_j is equal to β_j times the expected market excess return.

Assuming that the CAPM holds, we can hence describe the excess return of d assets at time $i \in \{1, \ldots, n\}$ via a regression model with one common covariate and without an intercept term, as specified in the following chapters.

Chapter 3

Bivariate regression normal copula model with a single common covariable

3.1 Introduction

In our application we have multivariate continuous response data available together with a common covariate. For this we want to construct univariate regression models for the margins and use a Gauss copula for modeling the dependence between the response variables. At first, we concentrate on the bivariate case. As financial application, one could possibly think of modeling the dependency of the returns of two asset classes given the market return in the CAPM framework described in Section 2.9. Note that the theoretical results of this chapter can also be derived by using the calculations in the chapter for three and more dimensions.

3.2 Model definition

Imagine we have a data set consisting of n pairs $(y_{i1}, y_{i2})'$ and n observations z_i, where $i \in \{1, \dots, n\}$. We want to describe the relationship between them using a two dimensional regression model on z_i with bivariate normal distributed errors to get random variables $(Y_{i1}, Y_{i2})'$ $(i = 1, \dots n)$, thus we can interpret the data $(y_{i1}, y_{i2})'$ as realizations of $(Y_{i1}, Y_{i2})'$ for all $i \in \{1, \dots, n\}$. Let $n \in \mathbb{N}$, $\boldsymbol{\beta} \in \mathbb{R}^2$, $\boldsymbol{\sigma}^2 \in (0, \infty)^2$ and $\rho \in (-1, 1)$.

Furthermore, let $\boldsymbol{\varepsilon}_i = \begin{pmatrix} \varepsilon_{i1} \\ \varepsilon_{i2} \end{pmatrix} \sim \mathcal{N}_2 \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} \right)$ i.i.d. $\forall i = 1, \dots, n$.

For given values $z_1 \in \mathbb{R}, \dots, z_n \in \mathbb{R}$ we define the two dimensional model by

$$Y_{ij} = z_i \beta_j + \sigma_j \varepsilon_{ij} \quad (i = 1, \dots, n)\ (j = 1, 2) \tag{3.1}$$

Using the notation $c(\cdot,\cdot|\rho)$ for the density of a bivariate Gauss copula with parameter ρ, we can also formulate the model as two dimensional pair-copula construction:

$$\begin{aligned} Y_{ij} &= z_i\beta_j + \sigma_j\varepsilon_{ij} && (i=1,\dots,n)\ (j=1,2) \\ p(\boldsymbol{\varepsilon}_i) &= c\left(\Phi(\varepsilon_{i1}),\Phi(\varepsilon_{i2})\middle|\rho\right)\varphi(\varepsilon_{i1})\varphi(\varepsilon_{i2}) && \text{i.i.d. } \forall i=1,\dots,n \end{aligned} \tag{3.2}$$

This means we have 5 unknown parameters in our model (3.2): The regression parameters β_1 and β_2, the residual variance parameters σ_1^2 and σ_2^2 and the residual correlation parameter ρ.

3.3 Prior choices

In a Bayesian setting, the unknown parameters β_j, σ_j^2 $(j=1,2)$ and ρ are not fixed values, but random variables, which means that we have to define prior distributions for each parameter.
Assume β_j, σ_j^2 $(j=1,2)$ and ρ are priorly independent and

$$\begin{aligned} \beta_j &\sim \mathcal{N}\left(0, s_j^2\right) && \text{independent} \quad \forall j=1,2 \\ \sigma_j^2 &\sim InverseGamma(a_j,b_j) && \text{independent} \quad \forall j=1,2 \\ \rho &\sim Uniform(-1,1) \end{aligned}$$

where $s_1^2, s_2^2 > 0$, $a_1, a_2 > 0$ and $b_1, b_2 > 0$ are parameters that can be chosen subject to the prior information.

Thus, the joint prior density $p(\boldsymbol{\beta}, \boldsymbol{\sigma}^2, \rho)$ is given as follows:

$$\begin{aligned} p(\boldsymbol{\beta},\boldsymbol{\sigma},\rho) &= \prod_{j=1}^{2} p(\beta_j)\cdot\prod_{j=1}^{2} p(\sigma_j^2)\cdot p(\rho) \\ p(\beta_j) &= \frac{1}{\sqrt{2\pi s_j^2}}\exp\left\{-\frac{\beta_j^2}{2s_j^2}\right\} && \forall j=1,2 \\ p(\sigma_j^2) &= \frac{b_j^{a_j}}{\Gamma(a_j)}(\sigma_j^2)^{-a_j-1}\exp\left\{-\frac{b_j}{\sigma_j^2}\right\} && \forall j=1,2 \\ p(\rho) &= \frac{1}{2} \end{aligned}$$

The assumption of prior independence of all parameters means that change of prior information on one parameter does not affect the prior distribution of the other parameters. The choice of a normal prior distribution for $\boldsymbol{\beta}$ is common for parameters that are proportional to the mean of the random variable of interest. Also the inverse gamma distribution is a suitable choice for positive parameters, so we use it as prior of $\boldsymbol{\sigma}^2$. For the last parameter ρ, we assume a non-informative prior, which is also consistent to our assumptions and results in higher dimensions. But any other prior for ρ can be easily integrated in the approach by multiplying it to the proportional expression of the full conditional density, which is derived later in this chapter.

3.4 Likelihood

We use the notation $\boldsymbol{Y}_i := (Y_{i1}, Y_{i2})'$ and as before $\boldsymbol{\beta} := (\beta_1, \beta_2)'$ and $\boldsymbol{\sigma}^2 := (\sigma_1^2, \sigma_2^2)'$. For the total response variables we write $\boldsymbol{Y} := (\boldsymbol{Y}'_1, \ldots, \boldsymbol{Y}'_n)'$.

From the model definition (3.2) we can directly derive the distribution of Y_i ($i \in \{1, \ldots, n\}$) for given values of $\boldsymbol{\beta}$, $\boldsymbol{\sigma}^2$, and ρ.

Lemma It holds $\boldsymbol{Y}_i \sim \mathcal{N}_2(\boldsymbol{\beta} z_i, \Sigma)$ for given values of $\boldsymbol{\beta}$, $\boldsymbol{\sigma}^2$, ρ and $i \in \{1, \ldots, n\}$, and the covariance matrix $\Sigma := \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}$ does not depend on i.

Proof. From (3.2) we know that $\boldsymbol{\varepsilon}_i \sim \mathcal{N}_2\left(0, \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}\right)$. From (2.2) we see that

$$\boldsymbol{Y}_i = \begin{pmatrix} z_i\beta_1 \\ z_i\beta_2 \end{pmatrix} + \begin{pmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{pmatrix} \boldsymbol{\varepsilon}_i \sim \mathcal{N}_2\Bigg(\begin{pmatrix} z_i\beta_1 \\ z_i\beta_2 \end{pmatrix}, \underbrace{\begin{pmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{pmatrix} \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} \begin{pmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{pmatrix}'}_{=\Sigma} \Bigg)$$

□

This means that for each $i \in \{1, \ldots, n\}$, the likelihood $f_i(\boldsymbol{Y}_i | \boldsymbol{\beta}, \boldsymbol{\sigma}^2, \rho)$ of $\boldsymbol{Y}_i$ corresponds to the density of a bivariate normal distribution, as defined in (2.1). As the error variables $\{\boldsymbol{\varepsilon}_i,\ i \in \{1, \ldots, n\}\}$ are independent, it follows from the model definition (3.2) that the response variables $\{\boldsymbol{Y}_i,\ i \in \{1, \ldots, n\}\}$ are independent for given values of $\boldsymbol{\beta}$, $\boldsymbol{\sigma}^2$ and ρ. Thus, the joint likelihood $f(\boldsymbol{Y} | \boldsymbol{\beta}, \boldsymbol{\sigma}^2, \rho)$ of all response variables is

$$\begin{aligned} f(\boldsymbol{Y} | \boldsymbol{\beta}, \boldsymbol{\sigma}^2, \rho) &= \prod_{i=1}^{n} f_i(\boldsymbol{Y}_i | \boldsymbol{\beta}, \boldsymbol{\sigma}, \rho) \\ &= \prod_{i=1}^{n} \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\Bigg\{ -\frac{1}{2(1-\rho^2)} \cdot \Bigg(\frac{(y_{i1} - z_i\beta_1)^2}{\sigma_1^2} + \frac{(y_{i2} - z_i\beta_2)^2}{\sigma_2^2} \\ &\qquad\qquad -2\rho\frac{(y_{i1} - z_i\beta_1)(y_{i2} - z_i\beta_2)}{\sigma_1\sigma_2} \Bigg) \Bigg\} \\ &= \frac{1}{(2\pi\sigma_1\sigma_2)^n (1-\rho^2)^{\frac{n}{2}}} \exp\Bigg\{ -\frac{1}{2(1-\rho^2)} \sum_{i=1}^{n} \Bigg(\frac{(y_{i1} - z_i\beta_1)^2}{\sigma_1^2} + \frac{(y_{i2} - z_i\beta_2)^2}{\sigma_2^2} \\ &\qquad\qquad -2\rho\frac{(y_{i1} - z_i\beta_1)(y_{i2} - z_i\beta_2)}{\sigma_1\sigma_2} \Bigg) \Bigg\} \end{aligned}$$

3.5 Full conditional distributions

Our aim is to construct an MCMC Algorithm to get samples from the joint posterior distribution of $(\boldsymbol{\beta}, \boldsymbol{\sigma}^2, \rho)$. To do this we follow the approach of a Gibbs sampler, that is sequentially updating the parameters $\boldsymbol{\beta}$, $\boldsymbol{\sigma}^2$ and ρ by using the full conditional distributions, i. e. the distribution of one parameter given all others and the data. However, we will see that only for $\boldsymbol{\beta}$ exists a known full conditional distribution, whereas we need Metropolis-Hastings steps for the other parameters. In this section, we derive the full conditional distribution of $\boldsymbol{\beta}$ and proportional expressions of the full conditional densities of ρ and $\phi_j^2 := \frac{1}{\sigma_j^2}$ $(j = 1, 2)$

3.5.1 Full conditional distribution of the regression parameter $\boldsymbol{\beta}$

We get a proportional expression of the full conditional density of $\boldsymbol{\beta}$ by multiplying the likelihood with the prior density of $\boldsymbol{\beta}$:

$$
\begin{aligned}
&p(\boldsymbol{\beta}|\boldsymbol{\sigma}^2, \rho, \boldsymbol{Y})p(\boldsymbol{\beta}) \propto f(\boldsymbol{Y}|\boldsymbol{\beta}, \boldsymbol{\sigma}^2, \rho) \\
&\propto \exp\left\{-\frac{1}{2}\left(\frac{\beta_1^2}{s_1^2} + \frac{\beta_2^2}{s_2^2}\right)\right\} \exp\Big\{ -\frac{1}{2(1-\rho^2)} \sum_{i=1}^{n} \Big(\frac{-2z_i y_{i1}\beta_1 + z_i^2\beta_1^2}{\sigma_1^2} + \frac{-2z_i y_{i2}\beta_2 + z_i^2\beta_2^2}{\sigma_2^2} \\
&\qquad\qquad - 2\rho\frac{z_i^2\beta_1\beta_2 - z_i y_{i1}\beta_2 - z_i y_{i2}\beta_1}{\sigma_1\sigma_2}\Big)\Big\} \\
&= \exp\Bigg\{ -\frac{1}{2}\Bigg(\frac{\sum_{i=1}^{n} z_i^2\beta_1^2}{\sigma_1^2(1-\rho^2)} + \frac{\beta_1^2}{s_1^2} + \frac{\sum_{i=1}^{n} z_i^2\beta_2^2}{\sigma_2^2(1-\rho^2)} + \frac{\beta_2^2}{s_2^2} - 2\frac{\rho\sum_{i=1}^{n} z_i^2\beta_1\beta_2}{\sigma_1\sigma_2(1-\rho^2)} \\
&\qquad - 2\frac{\sum_{i=1}^{n} y_{i1}z_i\beta_1}{\sigma_1^2(1-\rho^2)} + 2\frac{\rho\sum_{i=1}^{n} y_{i2}z_i\beta_1}{\sigma_1\sigma_2(1-\rho^2)} - 2\frac{\sum_{i=1}^{n} y_{i2}z_i\beta_2}{\sigma_2^2(1-\rho^2)} + 2\frac{\rho\sum_{i=1}^{n} y_{i1}z_i\beta_2}{\sigma_1\sigma_2(1-\rho^2)}\Bigg)\Bigg\}
\end{aligned}
$$

With $S_{zz} := \sum_{i=1}^{n} z_i^2$, $S_{zy_1} := \sum_{i=1}^{n} y_{i1}z_i$ and $S_{zy_2} := \sum_{i=1}^{n} y_{i2}z_i$ this term simplifies to

$$
\begin{aligned}
p(\boldsymbol{\beta}|\boldsymbol{\sigma}^2, \rho, \boldsymbol{Y}) \propto \exp\Bigg\{ -\frac{1}{2}\Bigg(&\frac{S_{zz}s_1^2 + \sigma_1^2(1-\rho^2)}{\sigma_1^2 s_1^2(1-\rho^2)}\beta_1^2 \quad + \frac{S_{zz}s_2^2 + \sigma_2^2(1-\rho^2)}{\sigma_2^2 s_2^2(1-\rho^2)}\beta_2^2 \\
&- 2\frac{\rho S_{zz}}{\sigma_1\sigma_2(1-\rho^2)}\beta_1\beta_2 \quad -2\frac{\sigma_2 S_{zy_1} - \rho\sigma_1 S_{zy_2}}{\sigma_1^2\sigma_2(1-\rho^2)}\beta_1 \\
&\qquad -2\frac{\sigma_1 S_{zy_2} - \rho\sigma_2 S_{zy_1}}{\sigma_1\sigma_2^2(1-\rho^2)}\beta_2\Bigg)\Bigg\} \qquad (3.3)
\end{aligned}
$$

We want to show that the full conditional density of $\boldsymbol{\beta}$ is the density of a bivariate normal distribution with mean vector $\boldsymbol{\mu} = \begin{pmatrix}\mu_1 \\ \mu_2\end{pmatrix}$ and covariance matrix $\begin{pmatrix}\tau_1^2 & \nu\tau_1\tau_2 \\ \nu\tau_1\tau_2 & \tau_2^2\end{pmatrix}$, where $\boldsymbol{\mu} \in \mathbb{R}^2$, $\tau_1 > 0$, $\tau_2 > 0$ and $\nu \in (-1, 1)$ have to be identified.

If this is the case, then

$$p(\boldsymbol{\beta}|\boldsymbol{\sigma}^2,\rho,\boldsymbol{Y}) \overset{!}{\propto} \exp\left\{ -\frac{1}{2(1-\nu^2)}\left(\frac{(\beta_1-\mu_1)^2}{\tau_1^2}+\frac{(\beta_2-\mu_2)^2}{\tau_2^2}-2\nu\frac{(\beta_1-\mu_1)(\beta_2-\mu_2)}{\tau_1\tau_2}\right)\right\}$$
$$\propto \exp\left\{ -\frac{1}{2}\left(\frac{\beta_1^2-2\mu_1\beta_1}{(1-\nu^2)\tau_1^2}+\frac{\beta_2^2-2\mu_2\beta_2}{(1-\nu^2)\tau_2^2}-2\nu\frac{\beta_1\beta_2-\mu_2\beta_1-\mu_1\beta_2}{(1-\nu^2)\tau_1\tau_2}\right)\right\}$$

Now we rearrange the expression above and get

$$p(\boldsymbol{\beta}|\boldsymbol{\sigma}^2,\rho,\boldsymbol{Y}) \overset{!}{\propto} \exp\left\{ -\frac{1}{2}\left(\frac{1}{(1-\nu^2)\tau_1^2}\beta_1^2+\frac{1}{(1-\nu^2)\tau_2^2}\beta_2^2-2\frac{\nu}{(1-\nu^2)\tau_1\tau_2}\beta_1\beta_2 \right.\right.$$
$$\left.\left. -2\frac{\tau_2\mu_1-\nu\tau_1\mu_2}{(1-\nu^2)\tau_1^2\tau_2}\beta_1-2\frac{\mu_2\tau_1-\nu\tau_2\mu_1}{(1-\nu^2)\tau_1\tau_2^2}\beta_2 \right)\right\} \tag{3.4}$$

By comparing expressions before β_1 in (3.3) and (3.4), it follows

$$\frac{1}{(1-\nu^2)\tau_1^2} \overset{!}{=} \frac{S_{zz}s_1^2+\sigma_1^2(1-\rho^2)}{\sigma_1^2 s_1^2(1-\rho^2)}$$
$$\Leftrightarrow \frac{1}{\sqrt{1-\nu^2}\tau_1} \overset{!}{=} \frac{\sqrt{S_{zz}s_1^2+\sigma_1^2(1-\rho^2)}}{\sigma_1 s_1\sqrt{1-\rho^2}} = \frac{v_1}{\sigma_1 s_1\sqrt{1-\rho^2}} \tag{3.5}$$

where $v_1 := \sqrt{S_{zz}s_1^2+\sigma_1^2(1-\rho^2)}$. In the same manner we see that

$$\frac{1}{\sqrt{1-\nu^2}\tau_2} \overset{!}{=} \frac{v_2}{\sigma_2 s_2\sqrt{1-\rho^2}} \tag{3.6}$$

with $v_2 := \sqrt{S_{zz}s_2^2+\sigma_2^2(1-\rho^2)}$.

If we now want to have equal factors before $\beta_1\beta_2$ in equations (3.3) and (3.4), it must hold

$$\frac{\nu}{(1-\nu^2)\tau_1\tau_2} \overset{!}{=} \frac{\rho S_{zz}}{\sigma_1\sigma_2(1-\rho^2)} \tag{3.7}$$

We can write the left hand side of (3.7) as $\nu\frac{1}{\sqrt{1-\nu^2}\tau_1}\frac{1}{\sqrt{1-\nu^2}\tau_2} \overset{(3.5)(3.6)}{=} \nu\frac{v_1v_2}{\sigma_1 s_1\sqrt{1-\rho^2}\sigma_2 s_2\sqrt{1-\rho^2}}$ which leads to

$$\nu = \frac{\rho S_{zz}s_1s_2}{v_1v_2} \tag{3.8}$$

Remark: Note that $\nu \in (-1,1)$ holds, since

$$v_1v_2 = \sqrt{S_{zz}s_1^2+\sigma_1^2(1-\rho^2)}\sqrt{S_{zz}s_2^2+\sigma_2^2(1-\rho^2)}$$
$$> \sqrt{S_{zz}s_1^2}\sqrt{S_{zz}s_2^2} > |\rho S_{zz}s_1s_2|$$
$$\Rightarrow |\nu| \overset{(3.8)}{=} \frac{|\rho S_{zz}s_1s_2|}{v_1v_2} < 1$$

From (3.8) it follows $1-\nu^2 = \frac{v_1^2 v_2^2 - \rho^2 S_{zz}^2 s_1^2 s_2^2}{v_1^2 v_2^2}$, so we get using (3.5)

$$\begin{aligned}\tau_1 &= \frac{\sigma_1 s_1 \sqrt{1-\rho^2}}{v_1\sqrt{1-\nu^2}} = \frac{\sigma_1 s_1 \sqrt{1-\rho^2} v_1 v_2}{v_1\sqrt{v_1^2 v_2^2 - \rho^2 S_{zz}^2 s_1^2 s_2^2}} \\ &= \frac{\sigma_1 s_1 \sqrt{1-\rho^2} v_2}{\sqrt{v_1^2 v_2^2 - \rho^2 S_{zz}^2 s_1^2 s_2^2}} = \frac{\sigma_1 s_1 \sqrt{1-\rho^2} v_2}{u} > 0 \end{aligned} \tag{3.9}$$

where we set $u := \sqrt{v_1^2 v_2^2 - \rho^2 S_{zz}^2 s_1^2 s_2^2}$.

Applying similar calculations to (3.6), we get

$$\tau_2 = \frac{\sigma_2 s_2 \sqrt{1-\rho^2} v_1}{u} > 0 \tag{3.10}$$

At last, we want to find an expression for the mean parameter $\boldsymbol{\mu} = (\mu_1, \mu_2)'$. For that purpose we again look at equations (3.3) and (3.4).
Comparing the factors before β_1, we see that

$$\begin{aligned} &\frac{\tau_2\mu_1 - \nu\tau_1\mu_2}{(1-\nu^2)\tau_1^2\tau_2} \overset{!}{=} \frac{\sigma_2 S_{zy_1} - \rho\sigma_1 S_{zy_2}}{\sigma_1^2\sigma_2(1-\rho^2)} \\ \Leftrightarrow \quad &\frac{\mu_1}{(1-\nu^2)\tau_1^2} = \frac{\sigma_2 S_{zy_1} - \rho\sigma_1 S_{zy_2}}{\sigma_1^2\sigma_2(1-\rho^2)} + \frac{\nu}{\tau_1}\frac{\mu_2}{(1-\nu^2)\tau_2} \end{aligned} \tag{3.11}$$

Similarly, we compare the factors before β_2:

$$\begin{aligned} &\frac{\tau_1\mu_2 - \nu\tau_2\mu_1}{(1-\nu^2)\tau_1\tau_2^2} \overset{!}{=} \frac{\sigma_1 S_{zy_2} - \rho\sigma_2 S_{zy_1}}{\sigma_1\sigma_2^2(1-\rho^2)} \\ \Leftrightarrow \quad &\frac{\mu_2}{(1-\nu^2)\tau_2} = \tau_2\left(\frac{\sigma_1 S_{zy_2} - \rho\sigma_2 S_{zy_1}}{\sigma_1\sigma_2^2(1-\rho^2)} + \frac{\nu}{\tau_2}\frac{\mu_1}{(1-\nu^2)\tau_1}\right) \end{aligned} \tag{3.12}$$

Setting (3.11) into (3.12) results in:

$$\begin{aligned} \frac{\mu_1}{(1-\nu^2)\tau_1^2} &= \frac{\sigma_2 S_{zy_1} - \rho\sigma_1 S_{zy_2}}{\sigma_1^2\sigma_2(1-\rho^2)} + \frac{\nu}{\tau_1}\tau_2\left(\frac{\sigma_1 S_{zy_2} - \rho\sigma_2 S_{zy_1}}{\sigma_1\sigma_2^2(1-\rho^2)} + \frac{\nu}{\tau_2}\frac{\mu_1}{(1-\nu^2)\tau_1}\right) \\ \Leftrightarrow \quad \frac{\mu_1}{\tau_1^2} &= \frac{\sigma_2 S_{zy_1} - \rho\sigma_1 S_{zy_2}}{\sigma_1^2\sigma_2(1-\rho^2)} + \frac{\nu\tau_2(\sigma_1 S_{zy_2} - \rho\sigma_2 S_{zy_1})}{\tau_1\sigma_1\sigma_2^2(1-\rho^2)} \end{aligned} \tag{3.13}$$

Now we solve (3.13) for μ_1:

$$\begin{aligned}\mu_1 &= \frac{\tau_1\left(\tau_1\sigma_2^2 S_{zy_1} - \rho\tau_1\sigma_1\sigma_2 S_{zy_2} + \nu\tau_2\sigma_1^2 S_{zy_2} - \rho\nu\tau_2\sigma_1\sigma_2 S_{zy_1}\right)}{\sigma_1^2\sigma_2^2(1-\rho^2)} \\ &= \frac{\left(\tau_1^2\sigma_2^2 - \rho\nu\tau_1\tau_2\sigma_1\sigma_2\right)S_{zy_1} + \left(\nu\tau_1\tau_2\sigma_1^2 - \rho\tau_1^2\sigma_1\sigma_2\right)S_{zy_2}}{\sigma_1^2\sigma_2^2(1-\rho^2)} \end{aligned} \tag{3.14}$$

Next, we want to eliminate τ_1, τ_2 and ν.
For that we use

$$\tau_1^2 \overset{(3.9)}{=} \frac{\sigma_1^2(1-\rho)^2 s_1^2 v_2^2}{u^2} \tag{3.15}$$

$$\text{and } \nu\tau_1\tau_2 \overset{(3.8)(3.9)(3.10)}{=} \frac{\rho S_{zz} s_1 s_2}{v_1 v_2} \frac{\sigma_1 s_1 \sqrt{1-\rho^2} v_2}{u} \frac{\sigma_2 s_2 \sqrt{1-\rho^2} v_1}{u} = \frac{\rho(1-\rho^2)\sigma_1\sigma_2 s_1^2 s_2^2 S_{zz}}{u^2} \tag{3.16}$$

This results in

$$\begin{aligned} \mu_1 &\overset{(3.14)(3.15)(3.16)}{=} \frac{(1-\rho^2)s_1^2\left(\left(\sigma_1^2\sigma_2^2 v_2^2 - \rho^2\sigma_1^2\sigma_2^2 s_2^2\right) S_{zy_1} + \left(\rho\sigma_1^3\sigma_2 s_2^2 S_{zz} - \rho\sigma_1^3\sigma_2 v_2^2\right) S_{zy_2}\right)}{u^2\sigma_1^2\sigma_2^2(1-\rho^2)} \\ &= \frac{s_1^2\left(\left(v_2^2 - \rho^2 s_2^2 S_{zz}\right)\sigma_2 S_{zy_1} + \left(s_2^2 S_{zz} - v_2^2\right)\rho\sigma_1 S_{zy_2}\right)}{\sigma_2 u^2} \end{aligned} \tag{3.17}$$

Finally, we use $v_2 - \rho^2 s_2^2 S_{zz} = S_{z^2y} s_2^2 + \sigma_2^2(1-\rho^2) - \rho^2 s_2^2 S_{z^2y} = (1-\rho^2)(S_{z^2y} s_2^2 + \sigma_2^2)$ and $s_2^2 S_{zz} - v_2^2 = -\sigma_2^2(1-\rho^2)$ to simplify (3.17) to

$$\mu_1 = \frac{(1-\rho^2)s_1^2}{u^2}\left((S_{zz}s_2^2 + \sigma_2^2)S_{zy_1} - \rho\sigma_1\sigma_2 S_{zy_2}\right) \tag{3.18}$$

Similarly, it follows that

$$\mu_2 = \frac{(1-\rho^2)s_2^2}{u^2}\left((S_{zz}s_1^2 + \sigma_1^2)S_{zy_2} - \rho\sigma_1\sigma_2 S_{zgy_1}\right) \tag{3.19}$$

Result: The full conditional density of $\boldsymbol{\beta}$ is a bivariate normal density whose parameters are given above.

3.5.2 Full conditional distribution of the correlation ρ

In the next part we determine an expression that is proportional to the full conditional density of ρ. As we have assumed a non-informative prior distribution for ρ, the full conditional density of ρ is proportional to the likelihood.

$$\begin{aligned} p(\rho|\boldsymbol{\beta},\boldsymbol{\sigma}^2,\boldsymbol{Y}) &\propto f(\boldsymbol{Y}|\boldsymbol{\beta},\boldsymbol{\sigma}^2,\rho)p(\rho) \propto f(\boldsymbol{Y}|\boldsymbol{\beta},\boldsymbol{\sigma}^2,\rho) \\ &\propto \frac{1}{(1-\rho^2)^{\frac{n}{2}}}\exp\left\{-\frac{1}{2(1-\rho^2)}\sum_{i=1}^{n}\left(\frac{(y_{i1}-z_i\beta_1)^2}{\sigma_1^2} + \frac{(y_{i2}-z_i\beta_2)^2}{\sigma_2^2}\right.\right. \\ &\qquad \left.\left. - 2\rho\frac{(y_{i1}-z_i\beta_1)(y_{i2}-z_i\beta_2)}{\sigma_1\sigma_2}\right)\right\} \end{aligned}$$

By defining $r_1^2 := \sum_{i=1}^{n}(y_{i1}-\beta_1 z_i)^2$, $r_2^2 := \sum_{i=1}^{n}(y_{i2}-\beta_2 z_i)^2$ and $r_{12} := \sum_{i=1}^{n}(y_{i1}-\beta_1 z_i)(y_{i2}-\beta_2 z_i)$ we can write the full conditional density of ρ as

$$p(\rho|\boldsymbol{\beta},\boldsymbol{\sigma}^2,\boldsymbol{Y}) \propto \frac{1}{(1-\rho^2)^{\frac{n}{2}}}\exp\left\{-\frac{1}{2(1-\rho^2)}\left(\frac{r_1^2}{\sigma_1^2} + \frac{r_2^2}{\sigma_2^2} - 2\rho\frac{r_{12}}{\sigma_1\sigma_2}\right)\right\}$$

We further use the abbreviations $a := \frac{r_1^2}{2\sigma_2^2} + \frac{r_2^2}{2\sigma_1^2}$ and $b := \frac{r_{12}}{\sigma_1\sigma_2}$, so that the full conditional density of ρ is proportional to a function $g_\rho(\rho, a, b)$ defined as follows:

$$g_\rho(\rho, a, b) := \frac{1}{(1-\rho^2)^{\frac{n}{2}}} \exp\left\{-\frac{a - b\rho}{1-\rho^2}\right\} \propto p(\rho|\boldsymbol{\beta}, \boldsymbol{\sigma}^2, \boldsymbol{Y}) \tag{3.20}$$

The function g_ρ is well defined and positive for all $\rho \in (-1, 1)$. Additionally, $a \geq b$ holds, since

$$\begin{aligned} a - b &= \frac{\sigma_2^2 \sum_{i=1}^n (y_{i1} - \beta_1 z_i)^2 + \sigma_1^2 \sum_{i=1}^n (y_{i2} - \beta_2 z_i)^2 - 2\sigma_1\sigma_2 \sum_{i=1}^n (y_{i1} - \beta_1 z_i)(y_{i2} - \beta_2 z_i)}{2\sigma_1^2\sigma_2^2} \\ &= \frac{\sum_{i=1}^n \sigma_2^2 (y_{i1} - \beta_1 z_i)^2 - 2\sigma_2\sigma_1 (y_{i1} - \beta_1 z_i)(y_{i2} - \beta_2 z_i) + \sigma_1^2 (y_{i2} - \beta_2 z_i)^2}{2\sigma_1^2\sigma_2^2} \\ &= \frac{\sum_{i=1}^n \left(\sigma_2 (y_{i1} - \beta_1 z_i) - \sigma_1 (y_{i2} - \beta_2 z_i)\right)^2}{2\sigma_1^2\sigma_2^2} \geq 0 \end{aligned}$$

However, a closed form for the integral $\int_{(-1,1)} g_\rho(\rho, a, b)\, \mathrm{d}\rho$ does not exist, which means that we know $p(\rho|\boldsymbol{\beta}, \boldsymbol{\sigma}^2, \boldsymbol{Y})$ only up to a proportional constant and therefore cannot use the Gibbs sampler to update ρ. Here, a Metropolis-Hastings step is necessary, and we will use $g_\rho(\rho, a, b)$ to calculate the acceptance probability. Due to numerical reasons, it is often better to calculate with the logarithm of $g_\rho(\rho, a, b)$, which is

$$\log(g_\rho(\rho, a, b)) = -\frac{n}{2} \log\left(1 - \rho^2\right) - \frac{a - b\rho}{1 - \rho^2} \tag{3.21}$$

3.5.3 Full conditional distribution of the error variances $\boldsymbol{\sigma}^2$

For the last parameter $\boldsymbol{\sigma}^2$, we do not directly calculate the full conditional density, but we look instead at $\phi_j^2 := \frac{1}{\sigma_j^2} \quad (j = 1, 2)$. This makes the calculations a bit easier. We know that from $\sigma_j^2 \sim InverseGamma(a_j, b_j)$ follows $\phi_j^2 \sim Gamma(a_j, b_j) (j = 1, 2)$. At first, we concentrate on ϕ_1^2.

$$\begin{aligned} &p(\phi_1^2|\phi_2^2, \boldsymbol{\beta}, \rho, \boldsymbol{Y}) \propto f(\boldsymbol{Y}|\boldsymbol{\beta}, \boldsymbol{\phi}^2, \rho) p(\phi_1^2) \\ &\propto (\phi_1^2)^{\frac{n}{2}} \exp\left\{-\frac{1}{2(1-\rho^2)} \sum_{i=1}^n \left(\phi_1^2 (y_{i1} - z_i\beta_1)^2 - 2\rho\phi_1\phi_2 (y_{i1} - z_i\beta_1)(y_{i2} - z_i\beta_2)\right)\right\} \\ &\quad \cdot \frac{b_1^{a_1}}{\Gamma(a_1)} (\phi_1^2)^{a_1 - 1} \exp\{-b_1\phi_1^2\} \\ &\propto (\phi_1^2)^{\frac{n-2}{2} + a_1} \exp\left\{-\frac{1}{2(1-\rho^2)} \left(\phi_1^2 \underbrace{\sum_{i=1}^n (y_{i1} - \beta_1 z_i)^2}_{=r_1^2} - 2\rho\phi_1\phi_2 \underbrace{\sum_{i=1}^n (y_{i1} - \beta_1 z_i)(y_{i2} - \beta_2)}_{=r_{12}}\right)\right\} \\ &\quad \cdot \exp\{-b_1\phi_1^2\} \end{aligned}$$

With the abbreviations $r_1^2 = \sum_{i=1}^{n}(y_{i1} - \beta_1 z_i)^2$ and $r_{12} = \sum_{i=1}^{n}(y_{i1} - \beta_1 z_i)(y_{i2} - \beta_2)$ that we have used before, we get

$$\begin{aligned} & p(\phi_1^2|\phi_2^2, \boldsymbol{\beta}, \rho, \boldsymbol{Y}) \\ & \quad \propto (\phi_1^2)^{\frac{n-2}{2}+a_1} \exp\left\{ -\frac{1}{2(1-\rho^2)}\left(\phi_1^2(r_1^2 + 2b_1(1-\rho^2)) - 2\rho\phi_1\phi_2 r_{12}\right)\right\} \\ & \quad = (\phi_1^2)^{\frac{n-2}{2}+a_1} \exp\left\{ -\frac{r_1^2 + 2b_1(1-\rho^2)}{2(1-\rho^2)}\left(\phi_1^2 - 2\sqrt{\phi_1^2}\frac{\rho\phi_2 r_{12}}{r_1^2 + 2b_1(1-\rho^2)}\right)\right\} \end{aligned} \tag{3.22}$$

By defining the function g_{ϕ_1} as

$$g_{\phi_1}(x, c, d) := x^{\frac{n-2}{2}+a_1} \exp\left\{-c\left(\sqrt{x} - d\right)^2\right\}$$

we can write the expression (3.22) in the form

$$p(\phi_1^2|\phi_2^2, \boldsymbol{\beta}, \rho, \boldsymbol{Y}) \propto g_{\phi_1}\left(\phi_1^2, \frac{r_1^2 + 2b_1(1-\rho^2)}{2(1-\rho^2)}, \frac{\rho\phi_2 r_{12}}{r_1^2 + 2b_1(1-\rho^2)}\right) = g_{\phi_1}(\phi_1^2, c_1, d_1)$$

where $c_1 := \frac{r_1^2 + 2b_1(1-\rho^2)}{2(1-\rho^2)}$ and $d_1 := \frac{\rho\phi_2 r_{12}}{r_1^2 + 2b_1(1-\rho^2)}$.

The logarithm of g_{ϕ_1} is given by

$$\log(g_{\phi_1}(x, c, d)) = \left(\frac{n-2}{2} + a_1\right)\log(x) - c\left(\sqrt{x} - d\right)^2 \tag{3.23}$$

For the parameter ϕ_2^2, we define the function g_{ϕ_2} as

$$g_{\phi_2}(x, c, d) := x^{\frac{n-2}{2}+a_2} \exp\left\{-c\left(\sqrt{x} - d\right)^2\right\}$$

The logarithm of g_{ϕ_2} is given by

$$\log(g_{\phi_2}(x, c, d)) = \left(\frac{n-2}{2} + a_2\right)\log(x) - c\left(\sqrt{x} - d\right)^2 \tag{3.24}$$

By performing similar calculations as before, one sees that

$$p(\phi_2^2|\phi_1^2, \boldsymbol{\beta}, \rho, \boldsymbol{Y}) \propto g_{\phi_2}\left(\phi_2^2, \frac{r_2^2 + 2b_2(1-\rho^2)}{2(1-\rho^2)}, \frac{\rho\phi_1 r_{12}}{r_2^2 + 2b_2(1-\rho^2)}\right) = g_{\phi_2}(\phi_2^2, c_2, d_2)$$

where $c_2 := \frac{r_2^2 + 2b_2(1-\rho^2)}{2(1-\rho^2)}$ and $d_2 := \frac{\rho\phi_2 r_{12}}{r_1^2 + 2b_1(1-\rho^2)}$.

3.6 The bivariate MCMC algorithm

Now that we have derived the full conditional density of $\boldsymbol{\beta}$ and proportional expressions for the full conditional densities of ρ, ϕ_1^2 and ϕ_2^2, we are able to build an MCMC algorithm to get samples from the posterior distribution. This algorithm constructs a hybrid chain consisting of a Gibbs sampler's update for $\boldsymbol{\beta}$ and Metropolis-Hastings steps for the other parameters. As mentioned before, it is not possible to construct a Gibbs Sampler for all parameters, since we completely know the full conditional density only for $\boldsymbol{\beta}$. On the other hand, a Metropolis-Hastings step for $\boldsymbol{\beta}$ would raise the problem of finding a two dimensional proposal density which does not lead to a high rate of rejections or would require to update the components β_1 and β_2 separately. The Gibbs sampler makes it possible to get new values for both components of $\boldsymbol{\beta}$ in one step without suffering any rejections. Therefore, the hybrid chain combines the advantage of the knowledge of the full conditional density of $\boldsymbol{\beta}$ with the possibility of updating the other parameters without completely knowing their full conditional density.

As proposal distribution for the Metropolis-Hastings steps, we take a random walk, i. e. a normal distribution whose mean is the old value for the parameter and whose variance is a predefined proposal variance $s^2_{prop,\cdot}$. This choice makes it easy both to sample from the proposal distribution and to calculate the acceptance probability. However, the random walk may propose negative values for ϕ_j, $j \in \{1, 2\}$ or inadmissable values for ρ. In those cases, the algorithm will of course reject the proposed values. This means that a high proposal variance may lead to a high rate of rejections.

Algorithm 3.1 MCMC Algorithm for two dimensions

INPUT data y_{ij} with $i = 1, \ldots, n$, $j = 1, 2$
  data z_i with $i = 1, \ldots, n$
  prior parameters s_1^2, s_2^2, a_1, b_1, a_2, b_2
  number of MCMC iterations m (resp. size of posterior distribution sample)
  proposal variances s^2_{prop,ϕ_1}, s^2_{prop,ϕ_2} and $s^2_{prop,\rho}$
  initial values $\boldsymbol{\beta}^{(0)}$, $\boldsymbol{\sigma}^{2(0)}$ and $\rho^{(0)}$
OUTPUT Samples $\boldsymbol{\beta}^{(1)}, \ldots, \boldsymbol{\beta}^{(m)}$, $\boldsymbol{\sigma}^{2(1)}, \ldots, \boldsymbol{\sigma}^{2(m)}$, $\rho^{(1)}, \ldots, \rho^{(m)}$

$\boldsymbol{\beta} := \boldsymbol{\beta}^{(0)}$, $\phi_1^2 := \frac{1}{\sigma_1^{2(0)}}$, $\phi_2^2 := \frac{1}{\sigma_2^{2(0)}}$, $\rho := \rho^{(0)}$
FOR $r := 1, \ldots, m$ **DO**
  $\boldsymbol{\beta}$-UPDATE:
  Draw a sample $\boldsymbol{x} = (x_1, x_2)'$ from $\mathcal{N}_2\left(\boldsymbol{\mu}, \begin{pmatrix} \tau_1^2 & \nu\tau_1\tau_2 \\ \nu\tau_1\tau_2 & \tau_2^2 \end{pmatrix}\right)$
    as defined in (3.8), (3.9), (3.10), (3.18) and (3.19)
  $\boldsymbol{\beta} := \boldsymbol{\beta}^{(r)} := \boldsymbol{x}$

Algorithm 3.2 MCMC Algorithm for two dimensions (continued)

$\boldsymbol{\sigma}^2$-UPDATE:
Draw a sample $\phi^2_{prop,1}$ from $\mathcal{N}(\phi^2_1, s^2_{prop,\phi_1})$
IF $\phi^2_{prop,1} > 0$ **THEN**
  Calculate logarithm of acceptance probability
    $l_{acc} := \max\left\{\log(g_{\phi_1}(\phi^2_{prop,1}, c_1, d_1)) - \log(g_{\phi_1}(\phi^2_1, c_1, d_1)), 0\right\}$
    using formula (3.23)
  Draw a sample x from $Uniform(0, 1)$
  IF $x \leq \exp\{l_{acc}\}$ **THEN**
    $\phi^2_1 := \phi^2_{prop,1}$
  END IF
END IF
Draw a sample $\phi^2_{prop,2}$ from $\mathcal{N}(\phi^2_2, s^2_{prop,\phi_2})$
IF $\phi^2_{prop,2} > 0$ **THEN**
  Calculate logarithm of acceptance probability
    $l_{acc} := \max\left\{\log(g_{\phi_2}(\phi^2_{prop,2}, c_2, d_2)) - \log(g_{\phi_2}(\phi^2_2, c_2, d_2)), 0\right\}$
    using formula (3.24)
  Draw a sample x from $Uniform(0, 1)$
  IF $x \leq \exp\{l_{acc}\}$ **THEN**
    $\phi^2_2 := \phi^2_{prop,2}$
  END IF
END IF
$\boldsymbol{\sigma}^{2(r)} := (\frac{1}{\phi^2_1}, \frac{1}{\phi^2_2})'$

ρ-UPDATE:
Draw a sample ρ_{prop} from $\mathcal{N}(\rho, s^2_{prop,\rho})$
IF $\rho_{prop} \in (-1, 1)$ **THEN**
  Calculate logarithm of acceptance probability
    $l_{acc} := \max\{\log(g_\rho(\rho_{prop}, a, b)) - \log(g_\rho(\rho, a, b)), 0\}$
    using formula (3.21)
   Draw a sample x from $Uniform(0, 1)$
  IF $x \leq \exp\{l_{acc}\}$ **THEN**
    $\rho := \rho_{prop}$
  END IF
END IF
$\rho^{(r)} := \rho$
END FOR

3.7 Small sample performance using the bivariate MCMC algorithm for Bayesian inference

In this section we test the derived algorithm on different scenarios and analyze the results. For that purpose we choose predefined values for the parameters $\boldsymbol{\beta}$, $\boldsymbol{\sigma}^2$ and ρ and create data based on these parameters. Then the algorithm uses this data to calculate estimates. We consider the correctness and convergence of the algorithm and look at the error behavior of the MCMC estimators. The two dimensional MCMC algorithm runs quite fast, which allows us to consider a lot of different scenarios and a high number of iterations.

3.7.1 Simulation setting

At first, we have to decide how much data we want to create. For every parameter constellation we consider two data sizes: $n = 1000$ and $n = 5000$. Whereas the first choice should be enough to get feasible estimates, we use the second one to consider if estimates improve when the data size is increased. Next, we have to specify the covariates z_i, which are independent of the parameter constellation. We choose the z_i's as a sequence from -1 to 1 with equal distance and length n, i. e. $z_1 := -1$, $z_n := 1$ and the other values are specified such that it holds:

$$z_i - z_{i-1} = \frac{2}{n-1} \quad \forall i \in \{2, \dots, n\}$$

So these values follow a line from -1 to 1, which means that the simulated data y_{ij} will follow a linear trend whose steepness depends on the value specified for β_j. For the $\beta' s$ we distinguish two situations: In the first part of the scenarios, we set $\beta_1 := \beta_2 := 0.5$, which means that y_{i1} and y_{i2} follow the same trend, whose steepness is lower than the one of the z_i's. For the second part of the scenarios, we change β_2 to 3, which means that trends of y_{i1} and y_{i2} are different: In the first case the trend of the z_i's is weakened, whereas the trend is strengthened in the second time series.

For the specification of σ_1^2 and σ_2^2, we look at the signal-to-noise ratio

$$SNR(Y_{ij}) := \frac{|E(Y_{ij})|}{\sqrt{Var(Y_{ij})}} \overset{(3.2)}{=} \frac{|z_i \beta_j|}{\sigma_j}, \quad i \in \{1, \dots, n\}, j \in \{1, 2\}$$

If the signal-to-noise ratio is greater than 2 for most of the data, then the signal dominates the noise, which should make estimation of the trend and correlation parameters easier. So it is of interest, which part of the data satisfies this relation, i. e. we are interested in the fraction

$$PSNR_j := \frac{\text{card}\{i \in \{1, \dots, n\} : SNR(Y_{ij}) > 2)\}}{n} = \frac{\text{card}\{i \in \{1, \dots, n\} : \frac{|z_i \beta_j|}{\sigma_j} > 2)\}}{n} \tag{3.25}$$

We will choose a $PSNR_j$ of 50% or of 80%, and we consider the case when these ratios are equal as well as the case when they are different. Since we have chosen $\beta_j > 0$ in all of

our scenarios and have used the same z_i's, we can infer σ_j^2 directly from β_j and $PSNR_j$. The resulting values of $\boldsymbol{\sigma}_j^2$ provided in the result tables are quite small, which is due to the choice of a low $\boldsymbol{\beta}$ and the data z_i, of which ca. 50% has an absolute value smaller 0.5.

For the important correlation parameter ρ, we consider no correlation, medium correlation and strong correlation. In Table 3.1, we list all 42 scenarios that we have considered. For each scenario, we perform 20 data replications, run the MCMC algorithm and analyze the results.

3.7.2 Further specification of the MCMC algorithm

There are still a few things to specify before we can run the MCMC algorithm. At first, we need initial values for the parameters $\boldsymbol{\beta}$, $\boldsymbol{\sigma}^2$ and ρ. For the marginal parameters, we take the marginal maximum likelihood estimators, which are

$$\beta_j^{(0)} := \frac{\sum_{i=1}^{n} z_i y_{ij}}{\sum_{i=1}^{n} z_i^2} = \frac{S_{zy_j}}{S_{zz}} \quad \text{and} \quad \sigma_j^{2(0)} := \frac{\sum_{i=1}^{n} (y_{ij} - z_i \beta_j^{(0)})^2}{n-1} = \frac{r_j^2}{n-1} \quad (j = 1, 2)$$

As initial value for ρ, we take the value that maximizes the conditional likelihood given $\boldsymbol{\beta} = \boldsymbol{\beta}^{(0)}$ and $\boldsymbol{\sigma}^2 = \boldsymbol{\sigma}^{2(0)}$. As the function g_ρ defined in (3.20) is proportional to the likelihood, we can use it for the maximization. However, this is equivalent to maximizing the logarithm of g_ρ that we derived in (3.21), but the latter is numerically more efficient.

This is why we set

$$\rho^{(0)} := \underset{\rho \in (-1,1)}{\arg\max} \{\log(g_\rho(\rho, a, b))\} \tag{3.26}$$

with $a := \frac{r_1^2}{2\sigma_2^{2(0)}} + \frac{r_2^2}{2\sigma_1^{2(0)}}$ and $b := \frac{r_{12}}{\sigma_1^{(0)}\sigma_2^{(0)}}$. To perform the maximization, we use the *R*-function *optimize()* which executes one dimensional optimizations.

Now that we have the initial values, we specify the prior parameters. At first, we have to specify the prior variances for β_1 and β_2. We decide to take an approximately non-informative prior distribution, which means that we take huge prior variances. Our choice is $s_1^2 := s_2^2 := 100000$. For σ_1^2 and σ_2^2, we follow Congdon (2003) on page 10 and set the parameters of the prior inverse gamma distribution to $a_1 := a_2 := 1$ and $b_1 := b_2 = 0.001$. From the preliminaries chapter, we know that this choice leads to a non-existing prior expectation for σ_1^2 and σ_2^2. As one can see, we have taken the same prior parameters for all components of $\boldsymbol{\beta}$ and $\boldsymbol{\sigma}^2$.

The proposal variance is determined by pilot runs for each parameter. This means that we take an initial value for the proposal variance of the considered parameter. Afterwards, we run the MCMC algorithm on the data with e.g. 2000 iterations and look at the acceptance rate of the parameter. If it is too small, e. g. less than 20%, we decrease the proposal

Scenario	$PSNR_1$	$PSNR_2$	β's	ρ	n
1	50%	50%	equal	0.0	1000
2	50%	50%	equal	0.0	5000
3	50%	50%	equal	0.5	1000
4	50%	50%	equal	0.5	5000
5	50%	50%	equal	0.8	1000
6	50%	50%	equal	0.8	5000
7	50%	80%	equal	0.0	1000
8	50%	80%	equal	0.0	5000
9	50%	80%	equal	0.5	1000
10	50%	80%	equal	0.5	5000
11	50%	80%	equal	0.8	1000
12	50%	80%	equal	0.8	5000
13	80%	80%	equal	0.0	1000
14	80%	80%	equal	0.0	5000
15	80%	80%	equal	0.5	1000
16	80%	80%	equal	0.5	5000
17	80%	80%	equal	0.8	1000
18	80%	80%	equal	0.8	5000
19	50%	50%	different	0.0	1000
20	50%	50%	different	0.0	5000
21	50%	50%	different	0.5	1000
22	50%	50%	different	0.5	5000
23	50%	50%	different	0.8	1000
24	50%	50%	different	0.8	5000
25	50%	80%	different	0.0	1000
26	50%	80%	different	0.0	5000
27	50%	80%	different	0.5	1000
28	50%	80%	different	0.5	5000
29	50%	80%	different	0.8	1000
30	50%	80%	different	0.8	5000
31	80%	50%	different	0.0	1000
32	80%	50%	different	0.0	5000
33	80%	50%	different	0.5	1000
34	80%	50%	different	0.5	5000
35	80%	50%	different	0.8	1000
36	80%	50%	different	0.8	5000
37	80%	80%	different	0.0	1000
38	80%	80%	different	0.0	5000
39	80%	80%	different	0.5	1000
40	80%	80%	different	0.5	5000
41	80%	80%	different	0.8	1000
42	80%	80%	different	0.8	5000

Table 3.1: Overview of all parameter constellations and data sizes used to test the algorithm

variance and repeat the procedure. The effect of this is that the proposed values will be closer to the old values, which should lead to less rejections and though to a higher acceptance rate. However, if the acceptance rate is too large, for instance greater than 80%, autocorrelations may get high, so it would be difficult to get approximately uncorrelated samples from the MCMC. Thus, we increase the proposal variance in that case and run the algorithm again, which has the opposite effect as described before. We carry out this procedure until we have proposal variances that lead to acceptance rates in a range of about 25% to 50%.

At last, we set the number of MCMC iterations to $m = 50000$, which is enough to get good estimates for the mode of each parameter and takes about 10 Minutes on my machine (Intel® Core™ 2 Duo CPU 2.2 GHz).

3.7.3 Results

As mentioned before, we simulate data 20 times and afterwards run our MCMC algorithm on each data set for each scenario specified in Table 3.1. When looking at the results, we see at first that the algorithm converges for every scenario and replication and that autocorrelations die down before lag 50. For each replication and each parameter $\theta \in \{\beta_1, \beta_2, \sigma_1^2, \sigma_2^2, \rho\}$, we then use every 50th iteration and a burn-in period of 5000 iterations to estimate the posterior density of θ with use of the *R*-function *density()*. This means that the density is estimated based on 900 approximately uncorrelated samples.

We show an example of the performance of the MCMC algorithm in Figure 3.1, in which we illustrate the results for one replication of scenario 21, i.e. with true parameter values $\beta_1 = 0.5$, $\beta_2 = 3$, $\sigma_1^2 = 0.0155$, $\sigma_2^2 = 0.563$ and $\rho = 0.5$. The left column of Figure 3.1 shows the parameter values of the MCMC algorithm for every 30th iteration. As one can see, the algorithm shows a good variation around the true values, which are marked as a horizontal line in the plots. Furthermore, the plots of the autocorrelation function located in the middle column indicate that autocorrelations get close to zero after a lag of approximately 30. This means that the MCMC mixes well in this example.

The plots in the right column show estimations of the posterior density based on our chain values. The two vertical lines mark the location of the mode and the true parameter value. Obviously we like to have the mode near to the true value, which is the case for most of the parameters. Only for ρ we see a greater distance between the two values, which is convenient to the corresponding trace plot in the left column, where most chain values lie below the line marking the true parameter value of 0.5. However, the estimated posterior density of the true parameter is high enough to be acceptable.

For each replication of each scenario, we get an estimator $\hat{\theta}_{mod}$ for the posterior mode θ_{mod} by taking the mode of the estimated density. We take this as estimate for each replication and calculate the mean of estimates $\bar{\theta}_{mod}$, the estimated bias, the relative bias and their standard errors. Recall that the estimated standard error of the mean is equal to that of the bias.

We present the results in Tables 3.2 to 3.8, whereas each table covers six scenarios with equal true values for $\boldsymbol{\beta}$ and $\boldsymbol{\sigma}^2$, but different values for n and ρ. The true values, estimates and standard errors are provided in the columns, where we used the following notations:

$$
\begin{aligned}
\text{Sc. \#} &: \text{Scenario index} \\
PSNR_j &: \text{Predifined } PSNR_j,\ j = 1, 2 \text{ for the scenario, see (3.25)} \\
n &: \text{length of simulated data set} \\
\theta &: \text{parameter of interest} \\
\theta_{true} &: \text{true value of } \theta \\
\bar{\theta}_{mod} &:= \frac{1}{20} \sum_{k=1}^{20} \hat{\theta}_{mod}^{(k)} \text{ , mean of estimated values } \hat{\theta}_{mod}^{(k)} \ (k = 1, \ldots, 20) \\
\widehat{b}(\hat{\theta}_{mod}) &:= \left(\frac{1}{20} \sum_{k=1}^{20} \hat{\theta}_{mod}^{(k)} \right) - \theta_{true} = \bar{\theta}_{mod} - \theta_{true} \text{ , estimated bias} \\
s(\bar{\theta}_{mod}) &:= \sqrt{\widehat{Var}\left(\bar{\theta}\right)} = \frac{s}{\sqrt{20}} \text{ , estimated standard error of } \bar{\theta}_{mod} \text{ and } \widehat{b}(\hat{\theta}_{mod}), \\
&\qquad \text{where } s := \sqrt{\frac{1}{19} \sum_{k=1}^{20} \left(\hat{\theta}_{mod}^{(k)} - \bar{\theta}_{mod} \right)} \\
\widehat{rb}(\hat{\theta}_{mod}) &:= \frac{\bar{\theta}_{mod} - \theta_{true}}{\theta_{true}} = \frac{\widehat{b}(\hat{\theta}_{mod})}{\theta_{true}} \text{ , estimated relative bias} \\
s_{rb}(\hat{\theta}_{mod}) &:= \sqrt{\widehat{Var}\left(\frac{\bar{\theta}_{mod} - \theta_{true}}{\theta_{true}} \right)} = \sqrt{\frac{s^2}{20} \frac{1}{\theta_{true}^2}} = \frac{s(\bar{\theta}_{mod})}{\theta_{true}}, \\
&\qquad \text{estimated standard error of } \widehat{rb}(\hat{\theta}_{mod})
\end{aligned}
$$

For each table, the bold values in the $\widehat{b}(\hat{\theta}_{mod})$ column show the maximum absolute values for the same $\theta \in \{\beta_1, \beta_2, \sigma_1^2, \sigma_2^2\}$ over all scenarios mentioned in the table. The largest absolute value of the estimated bias of ρ is not highlighted, since the true value of ρ changes three times, and therefore the bias estimates from different scenarios listed in the same table are not always comparable. For the $\widehat{rb}(\hat{\theta}_{mod})$ column, we instead mark the highest absolute value within one scenario in bold type. This makes sense, since the relative bias (in contrast to the bias itself) makes it possible to compare the accuracy of estimates between parameters even when their true values are different.

We do not calculate the estimated mean squared error because a plausible value for it is hardly achieved with only 20 replications. Furthermore, there are some scenarios where the true value for the parameter ρ is 0. Here, the relative bias obviously makes no sense resp. would be ∞. In those cases, we use the notation "–".

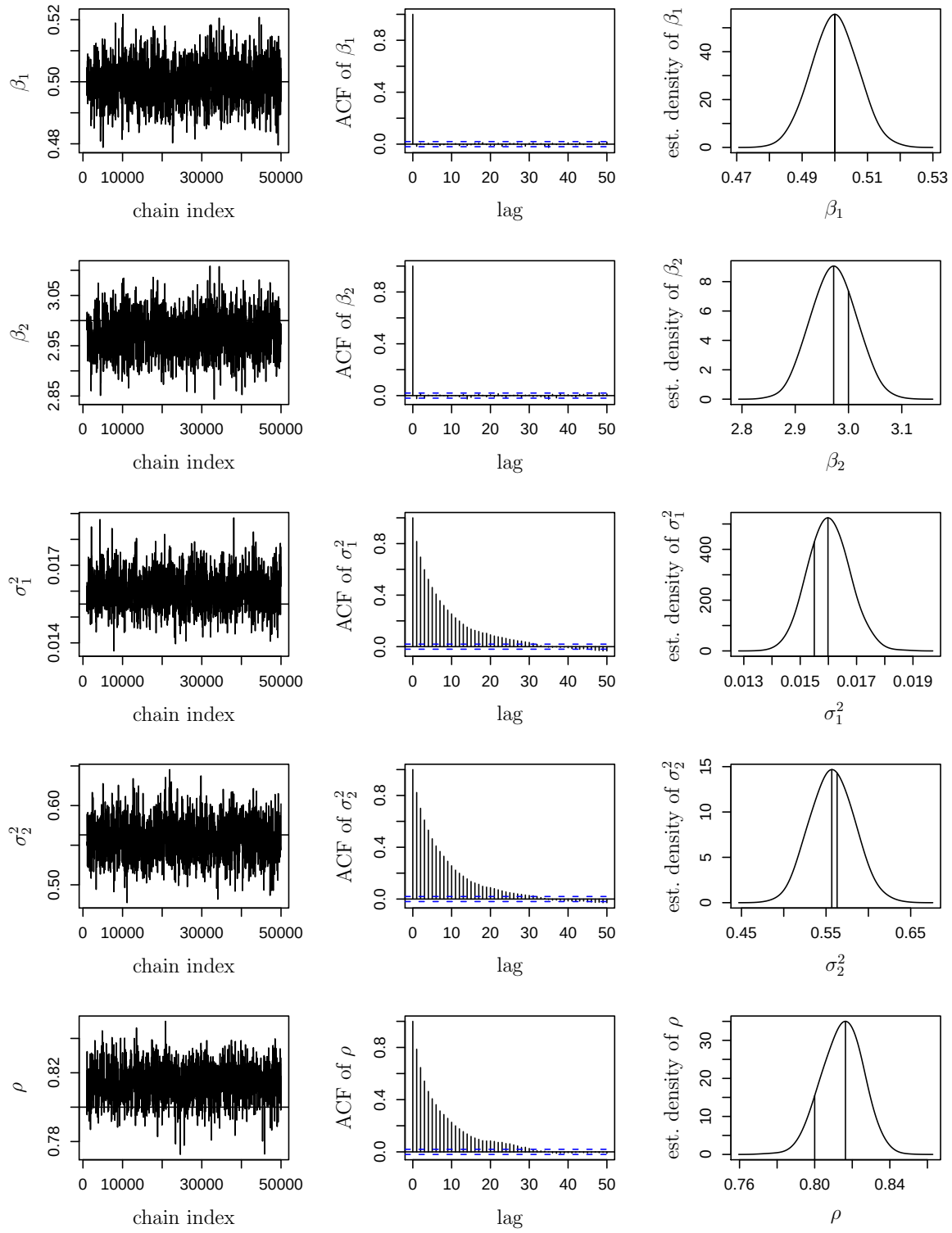

Figure 3.1: Trace plots, autocorrelation plots and estimated density for one replication of Scenario 23

Sc. #	$PSNR_1$	$PSNR_2$	n	θ	θ_{true}	$\bar{\theta}_{mod}$	$10^2 \cdot \widehat{b}(\widehat{\theta}_{mod})$	$10^2 \cdot s(\bar{\theta}_{mod})$	$10^2 \cdot \widehat{rb}(\widehat{\theta}_{mod})$	$10^2 \cdot s_{rb}(\widehat{\theta}_{mod})$
1	0.5	0.5	1000	β_1	0.5000	0.4981	**−0.191**	0.139	−0.382	0.279
				β_2	0.5000	0.5017	**0.167**	0.176	0.333	0.353
				σ_1^2	0.0155	0.0156	0.014	0.011	0.929	0.720
				σ_2^2	0.0155	0.0153	−0.020	0.013	**−1.287**	0.835
				ρ	0.0000	−0.0021	−0.214	0.585	–	–
2	0.5	0.5	5000	β_1	0.5000	0.5009	0.087	0.055	0.173	0.109
				β_2	0.5000	0.4997	−0.032	0.072	−0.065	0.144
				σ_1^2	0.0155	0.0154	−0.005	0.004	−0.326	0.288
				σ_2^2	0.0155	0.0153	−0.019	0.006	**−1.218**	0.418
				ρ	0.0000	0.0019	0.188	0.207	–	–
3	0.5	0.5	1000	β_1	0.5000	0.5003	0.033	0.161	0.067	0.322
				β_2	0.5000	0.5001	0.011	0.147	0.023	0.295
				σ_1^2	0.0155	0.0152	**−0.027**	0.015	**−1.750**	0.978
				σ_2^2	0.0155	0.0153	**−0.022**	0.016	−1.393	1.038
				ρ	0.5000	0.5008	0.083	0.582	0.167	1.163
4	0.5	0.5	5000	β_1	0.5000	0.4989	−0.106	0.058	−0.211	0.115
				β_2	0.5000	0.4997	−0.025	0.072	−0.051	0.143
				σ_1^2	0.0155	0.0154	−0.015	0.005	**−0.947**	0.311
				σ_2^2	0.0155	0.0154	−0.010	0.007	−0.633	0.450
				ρ	0.5000	0.4969	−0.308	0.241	−0.616	0.483
5	0.5	0.5	1000	β_1	0.5000	0.4991	−0.088	0.137	−0.176	0.274
				β_2	0.5000	0.4983	−0.165	0.088	−0.331	0.176
				σ_1^2	0.0155	0.0154	−0.006	0.014	**−0.365**	0.885
				σ_2^2	0.0155	0.0155	−0.001	0.017	−0.060	1.077
				ρ	0.8000	0.7988	−0.118	0.202	−0.147	0.253
6	0.5	0.5	5000	β_1	0.5000	0.5003	0.026	0.060	0.052	0.119
				β_2	0.5000	0.5002	0.019	0.064	0.037	0.127
				σ_1^2	0.0155	0.0155	−0.005	0.007	**−0.305**	0.444
				σ_2^2	0.0155	0.0155	−0.002	0.006	−0.148	0.364
				ρ	0.8000	0.8010	0.103	0.123	0.129	0.154

Table 3.2: Mean of estimates, estimated bias, relative bias and their estimated standard errors for different data size and correlations with $\beta_1 = \beta_2$ and low $PSNR_1 = PSNR_2$

Scenarios with equal β's and low $PSNR_1 = PSNR_2$

We start with analyzing the scenarios where $\beta_1 = \beta_2$ holds. In Table 3.2 we see the results for the scenarios with $PSNR_1 = PSNR_2 = 0.5$. The largest deviation between estimates and true values for both β_1 and β_2 can be found in the first scenario. with no correlation and the smaller data size $n = 1000$. Here, the bias is about -0.19% for β_1 and 0.17% for β_2. Since the true parameter values of σ_1^2 and σ_2^2 are very small, it does not surprise that this also holds for the estimated bias, whose largest absolute values can be found in scenario 3. With only one exception (σ_1^2 in scenarios 5 and 6), the estimate for the residual variance parameters σ_j^2 gets closer to the true value when n is increased.

To compare the estimation between different parameters, we look at the relative bias. We observe that the largest relative deviation from the true value can always be observed for a σ_j^2 estimate. The highest relative deviation is that of the σ_1^2 estimate in scenario 3 with about 1.75%. The estimated bias for ρ ranges from -0.31% to -0.10%.

The standard errors of the relative bias $s_{rb}(\widehat{\theta}_{mod})$ are all smaller than 1.2% and decreasing when n goes from 1000 to 5000. We observe the decreasing standard error when increasing n in all 42 scenarios so we will not mention this again when analyzing Tables 3.3 to 3.8.

Sc. #	$PSNR_1$	$PSNR_2$	n	θ	θ_{true}	$\bar{\theta}_{mod}$	$10^2 \cdot \widehat{b}(\widehat{\theta}_{mod})$	$10^2 \cdot s(\bar{\theta}_{mod})$	$10^2 \cdot \widehat{rb}(\widehat{\theta}_{mod})$	$10^2 \cdot s_{rb}(\widehat{\theta}_{mod})$
7	0.5	0.8	1000	β_1	0.5000	0.4989	−0.109	0.154	−0.219	0.309
				β_2	0.5000	0.5000	0.003	0.075	0.005	0.150
				σ_1^2	0.0155	0.0157	**0.017**	0.017	**1.103**	1.108
				σ_2^2	0.0025	0.0025	0.002	0.003	0.900	1.027
				ρ	0.0000	−0.0102	−1.018	0.809	–	–
8	0.5	0.8	5000	β_1	0.5000	0.4997	−0.035	0.057	−0.069	0.114
				β_2	0.5000	0.5006	**0.064**	0.026	0.129	0.051
				σ_1^2	0.0155	0.0155	−0.004	0.006	**−0.277**	0.407
				σ_2^2	0.0025	0.0025	0.000	0.001	0.033	0.519
				ρ	0.0000	0.0031	0.307	0.474	–	–
9	0.5	0.8	1000	β_1	0.5000	0.5000	0.003	0.174	0.006	0.349
				β_2	0.5000	0.5002	0.020	0.050	0.040	0.100
				σ_1^2	0.0155	0.0154	−0.009	0.014	−0.554	0.931
				σ_2^2	0.0025	0.0025	−0.002	0.003	**−0.752**	1.099
				ρ	0.5000	0.4985	−0.153	0.455	−0.307	0.910
10	0.5	0.8	5000	β_1	0.5000	0.4985	**−0.146**	0.089	−0.292	0.178
				β_2	0.5000	0.4996	−0.040	0.021	−0.080	0.042
				σ_1^2	0.0155	0.0156	0.010	0.006	0.670	0.381
				σ_2^2	0.0025	0.0025	**0.003**	0.001	**1.088**	0.585
				ρ	0.5000	0.5025	0.249	0.190	0.498	0.380
11	0.5	0.8	1000	β_1	0.5000	0.4988	−0.120	0.149	−0.241	0.297
				β_2	0.5000	0.4997	−0.032	0.059	−0.064	0.119
				σ_1^2	0.0155	0.0154	−0.013	0.016	**−0.821**	1.003
				σ_2^2	0.0025	0.0025	−0.001	0.003	−0.307	1.036
				ρ	0.8000	0.7966	−0.342	0.241	−0.428	0.301
12	0.5	0.8	5000	β_1	0.5000	0.5002	0.022	0.067	0.045	0.133
				β_2	0.5000	0.5003	0.029	0.024	0.057	0.049
				σ_1^2	0.0155	0.0154	−0.005	0.006	**−0.355**	0.387
				σ_2^2	0.0025	0.0025	0.000	0.001	0.078	0.401
				ρ	0.8000	0.8012	0.119	0.079	0.149	0.099

Table 3.3: Mean of estimates, estimated bias, relative bias and their estimated standard errors for different data size and correlations with $\beta_1 = \beta_2$ and $PSNR_1 < PSNR_2$

Scenarios with equal β's and different $PSNR_j$

In the next six scenarios, we increase $PSNR_2$ to 80%, while all other true values stay the same. The results are provided in Table 3.3.

The maximum absolute value of the estimated bias of β_1 can be found in scenario 10, that of β_2 in scenario 8, both with $n = 5000$. For the σ_j^2 parameters we find the according maximums in scenarios 7 and 10. As expected, this value is much larger for σ_1^2, since the true value of σ_2^2 has decreased due to the higher $PSNR_2$. For the scenario with high correlation, the deviation from the true value is in general smaller than in those with no or medium correlation.

When comparing the relative bias of all parameters within a scenario, we see that the largest absolute values always belongs to a σ_j^2 parameter. The largest of them is that of σ_1^2 in scenario 7, with no correlation and the smaller data size. For $\rho \neq 0$ we observe estimated relative biases between -0.43% and 0.50%. The biases in the uncorrelated settings (scenarios 7 and 8) are -1.02% and 0.31%.

Sc. #	$PSNR_1$	$PSNR_2$	n	θ	θ_{true}	$\bar{\theta}_{mod}$	$10^2 \cdot \widehat{b}(\widehat{\theta}_{mod})$	$10^2 \cdot s(\bar{\theta}_{mod})$	$10^2 \cdot \widehat{rb}(\widehat{\theta}_{mod})$	$10^2 \cdot s_{rb}(\widehat{\theta}_{mod})$
13	0.8	0.8	1000	β_1	0.5000	0.5006	0.061	0.045	0.122	0.090
				β_2	0.5000	0.4999	−0.006	0.052	−0.012	0.105
				σ_1^2	0.0025	0.0025	**0.003**	0.003	**1.258**	1.105
				σ_2^2	0.0025	0.0025	−0.002	0.003	−0.930	1.149
				ρ	0.0000	−0.0003	−0.027	0.531	–	–
14	0.8	0.8	5000	β_1	0.5000	0.4999	−0.010	0.028	−0.019	0.056
				β_2	0.5000	0.4998	−0.017	0.029	−0.034	0.057
				σ_1^2	0.0025	0.0025	−0.000	0.001	−0.196	0.387
				σ_2^2	0.0025	0.0025	0.001	0.001	**0.495**	0.356
				ρ	0.0000	−0.0030	−0.301	0.318	–	–
15	0.8	0.8	1000	β_1	0.5000	0.4993	−0.069	0.058	−0.137	0.115
				β_2	0.5000	0.4998	−0.018	0.066	−0.037	0.131
				σ_1^2	0.0025	0.0025	0.000	0.003	0.026	1.020
				σ_2^2	0.0025	0.0025	0.002	0.002	**0.807**	0.990
				ρ	0.5000	0.5005	0.053	0.584	0.105	1.168
16	0.8	0.8	5000	β_1	0.5000	0.4998	−0.017	0.024	−0.034	0.047
				β_2	0.5000	0.4999	−0.010	0.025	−0.019	0.051
				σ_1^2	0.0025	0.0025	−0.000	0.001	−0.109	0.494
				σ_2^2	0.0025	0.0025	−0.001	0.001	**−0.425**	0.340
				ρ	0.5000	0.5016	0.161	0.169	0.322	0.338
17	0.8	0.8	1000	β_1	0.5000	0.4982	**−0.182**	0.036	−0.363	0.073
				β_2	0.5000	0.4988	**−0.118**	0.030	−0.236	0.061
				σ_1^2	0.0025	0.0025	0.001	0.002	0.570	0.870
				σ_2^2	0.0025	0.0025	**0.003**	0.002	**1.249**	0.966
				ρ	0.8000	0.8025	0.255	0.253	0.318	0.317
18	0.8	0.8	5000	β_1	0.5000	0.4998	−0.018	0.031	−0.037	0.062
				β_2	0.5000	0.4999	−0.013	0.032	−0.027	0.064
				σ_1^2	0.0025	0.0025	0.001	0.001	**0.354**	0.428
				σ_2^2	0.0025	0.0025	0.000	0.001	0.116	0.383
				ρ	0.8000	0.7992	−0.081	0.089	−0.101	0.111

Table 3.4: Mean of estimates, estimated bias, relative bias and their estimated standard errors for different data size and correlations with $\beta_1 = \beta_2$ and high $PSNR_1 = PSNR_2$

Scenarios with equal β's and high $PSNR_1 = PSNR_2$

Table 3.4 provides the results when $PSNR_1$ is changed to 80%, such that $PSNR_1$ and $PSNR_2$ are equal again.

Here, we find the maximum absolute value of the bias for the β_j parameters in the scenario with high correlation and $n = 1000$, namely 0.18% for β_1 and 0.12% for β_2. The absolute values of the estimated biases for σ_1^2 and σ_2^2 are now both very small with a maximum absolute value of 0.03% for σ_1^2 in scenario 13 and 0.03% for σ_2^2 in scenario 17. The appropriate relative biases are about 1.25%.

All corresponding relative biases for $\rho \neq 0$ are smaller than 0.33%. Again, the largest absolute value of the relative bias within a scenario is always that of σ_1^2 or σ_2^2. We frequently observe that the relative bias gets closer to 0 when n is increased, this especially holds for the scenarios with medium or high correlation.

Sc. #	$PSNR_1$	$PSNR_2$	n	θ	θ_{true}	$\bar{\theta}_{mod}$	$10^2 \cdot \widehat{b}(\widehat{\theta}_{mod})$	$10^2 \cdot s(\bar{\theta}_{mod})$	$10^2 \cdot \widehat{rb}(\widehat{\theta}_{mod})$	$10^2 \cdot s_{rb}(\widehat{\theta}_{mod})$
19	0.5	0.5	1000	β_1	0.5000	0.4965	**−0.345**	0.157	−0.690	0.315
				β_2	3.0000	3.0069	**0.688**	0.842	0.229	0.281
				σ_1^2	0.0155	0.0154	−0.011	0.020	**−0.726**	1.274
				σ_2^2	0.5630	0.5653	0.234	0.487	0.416	0.864
				ρ	0.0000	−0.0092	−0.917	0.785	–	–
20	0.5	0.5	5000	β_1	0.5000	0.4995	−0.051	0.072	−0.102	0.145
				β_2	3.0000	3.0004	0.040	0.399	0.013	0.133
				σ_1^2	0.0155	0.0155	−0.005	0.006	−0.298	0.378
				σ_2^2	0.5630	0.5651	0.208	0.278	**0.370**	0.494
				ρ	0.0000	0.0005	0.054	0.354	–	–
21	0.5	0.5	1000	β_1	0.5000	0.4986	−0.143	0.149	−0.287	0.299
				β_2	3.0000	2.9980	−0.202	1.044	−0.067	0.348
				σ_1^2	0.0155	0.0154	−0.011	0.021	**−0.716**	1.359
				σ_2^2	0.5630	0.5624	−0.059	0.493	−0.105	0.875
				ρ	0.5000	0.4990	−0.097	0.499	−0.195	0.997
22	0.5	0.5	5000	β_1	0.5000	0.5004	0.037	0.062	0.075	0.124
				β_2	3.0000	2.9973	−0.268	0.536	−0.089	0.179
				σ_1^2	0.0155	0.0156	**0.014**	0.007	0.927	0.419
				σ_2^2	0.5630	0.5668	**0.384**	0.238	0.682	0.423
				ρ	0.5000	0.5073	0.728	0.284	**1.457**	0.567
23	0.5	0.5	1000	β_1	0.5000	0.4986	−0.143	0.135	−0.286	0.271
				β_2	3.0000	2.9933	−0.669	0.728	−0.223	0.243
				σ_1^2	0.0155	0.0154	−0.010	0.020	**−0.623**	1.280
				σ_2^2	0.5630	0.5605	−0.251	0.637	−0.446	1.131
				ρ	0.8000	0.7982	−0.184	0.374	−0.230	0.468
24	0.5	0.5	5000	β_1	0.5000	0.4989	−0.111	0.079	−0.222	0.157
				β_2	3.0000	2.9978	−0.216	0.501	−0.072	0.167
				σ_1^2	0.0155	0.0154	−0.006	0.009	**−0.384**	0.563
				σ_2^2	0.5630	0.5612	−0.176	0.255	−0.313	0.454
				ρ	0.8000	0.7987	−0.127	0.084	−0.159	0.105

Table 3.5: Mean of estimates, estimated bias, relative bias and their estimated standard errors for different data size and correlations with $\beta_1 < \beta_2$ and low $PSNR_1 = PSNR_2$

Scenarios with different β's and low $PSNR_1 = PSNR_2$

So far, we looked at situations where the true β_j parameters are the same. Now we increase β_2 from 0.5 to 3 and analyze the results, which are provided in Tables 3.5 to 3.8. The first setting shown in Table 3.5 is comparable to that in Table 3.2, except the fact that we have different β_j parameters.

We find the maximum absolute value for the bias of both β_1 and β_2 in scenario 19. The corresponding maximum values for the residual variance parameters are located in scenario 22, where this value is 0.14% for σ_1^2 and 0.38% for σ_2^2. The larger value of σ_2^2 can be explained by the fact that the true value of β_2 is larger than that of β_1 and therefore the true value of σ_2^2 derived from the predefined $PSNR_2$ is larger.

For the relative biases, we observe that their absolute values drop for all parameters when n is increased, except for scenarios 21 and 22. Similarly as we have observed it before, the largest absolute value of $\widehat{rb}(\widehat{\theta}_{mod})$ within a scenario belongs to a residual variance parameter, except for scenario 22, where that of ρ is greater than that of the other parameters. Here the estimated relative bias is 1.46%, which is also the maximum absolute value of all relative bias estimates in the table.

Sc. #	$PSNR_1$	$PSNR_2$	n	θ	θ_{true}	$\bar{\theta}_{mod}$	$10^2 \cdot \widehat{b}(\widehat{\theta}_{mod})$	$10^2 \cdot s(\bar{\theta}_{mod})$	$10^2 \cdot \widehat{rb}(\widehat{\theta}_{mod})$	$10^2 \cdot s_{rb}(\widehat{\theta}_{mod})$
25	0.5	0.8	1000	β_1	0.5000	0.5005	0.047	0.162	0.093	0.324
				β_2	3.0000	3.0047	**0.472**	0.483	0.157	0.161
				σ_1^2	0.0155	0.0157	**0.022**	0.017	**1.409**	1.105
				σ_2^2	0.0900	0.0901	0.011	0.079	0.124	0.874
				ρ	0.0000	−0.0055	−0.548	0.671	–	–
26	0.5	0.8	5000	β_1	0.5000	0.5008	0.078	0.061	0.156	0.121
				β_2	3.0000	2.9965	−0.346	0.126	−0.115	0.042
				σ_1^2	0.0155	0.0154	−0.008	0.005	**−0.518**	0.339
				σ_2^2	0.0900	0.0901	0.007	0.033	0.076	0.369
				ρ	0.0000	−0.0040	−0.405	0.309	–	–
27	0.5	0.8	1000	β_1	0.5000	0.5011	0.106	0.224	0.212	0.448
				β_2	3.0000	3.0002	0.021	0.469	0.007	0.156
				σ_1^2	0.0155	0.0154	−0.007	0.015	−0.457	0.969
				σ_2^2	0.0900	0.0895	**−0.052**	0.069	−0.579	0.765
				ρ	0.5000	0.4969	−0.314	0.423	**−0.628**	0.845
28	0.5	0.8	5000	β_1	0.5000	0.5003	0.032	0.072	0.064	0.144
				β_2	3.0000	3.0010	0.098	0.135	0.033	0.045
				σ_1^2	0.0155	0.0155	−0.003	0.008	−0.166	0.484
				σ_2^2	0.0900	0.0901	0.010	0.046	0.115	0.515
				ρ	0.5000	0.4990	−0.103	0.242	**−0.206**	0.484
29	0.5	0.8	1000	β_1	0.5000	0.4968	**−0.322**	0.131	**−0.644**	0.261
				β_2	3.0000	2.9967	−0.333	0.338	−0.111	0.113
				σ_1^2	0.0155	0.0154	−0.010	0.015	−0.639	0.951
				σ_2^2	0.0900	0.0904	0.044	0.092	0.490	1.018
				ρ	0.8000	0.7960	−0.404	0.212	−0.505	0.266
30	0.5	0.8	5000	β_1	0.5000	0.5007	0.066	0.059	0.131	0.118
				β_2	3.0000	2.9996	−0.044	0.182	−0.015	0.061
				σ_1^2	0.0155	0.0154	−0.014	0.006	**−0.899**	0.371
				σ_2^2	0.0900	0.0898	−0.019	0.041	−0.216	0.460
				ρ	0.8000	0.7992	−0.083	0.087	−0.103	0.109

Table 3.6: Mean of estimates, estimated bias, relative bias and their estimated standard errors for different data size and correlations with $\beta_1 < \beta_2$ and $PSNR_1 < PSNR_2$

Besides the logical consequences from the change of the true values of β_2 and σ_2^2 on the bias, we see no systematic change when we compare the results to those of scenarios 1 to 6. But one can observe that the standard errors of the estimated relative bias stay similar.

Scenarios with different β's and different $PSNR_j$

For the scenarios 25 to 30, we change the value of $PSNR_2$ to 80%, so the settings belonging to Table 3.6 can be compared to the scenarios mentioned in Table 3.3, besides the different predefinition of $\boldsymbol{\beta}$.

The maximum absolute value of the estimated bias of β_1 is located in scenario 29, that of β_2 in scenario 25. For σ_1^2 and σ_2^2 we see that the highest absolute values of the bias are those in scenarios 25 and 27. This means that all four maximum values belong to scenarios where $n = 1000$ holds.

When we compare the relative deviation from the true values, we see that the absolute value of the estimated relative bias within a scenario ranges between 0.21% and 1.41%, and drops for most scenarios when n is increased.

Sc. #	$PSNR_1$	$PSNR_2$	n	θ	θ_{true}	$\bar{\theta}_{mod}$	$10^2 \cdot \widehat{b}(\widehat{\theta}_{mod})$	$10^2 \cdot s(\bar{\theta}_{mod})$	$10^2 \cdot \widehat{rb}(\widehat{\theta}_{mod})$	$10^2 \cdot s_{rb}(\widehat{\theta}_{mod})$
31	0.8	0.5	1000	β_1	0.5000	0.4996	−0.037	0.074	−0.075	0.148
				β_2	3.0000	2.9992	−0.076	1.014	−0.025	0.338
				σ_1^2	0.0025	0.0025	**−0.002**	0.003	**−0.878**	1.137
				σ_2^2	0.5630	0.5675	0.452	0.522	0.804	0.927
				ρ	0.0000	0.0025	0.248	0.812	–	–
32	0.8	0.5	5000	β_1	0.5000	0.5000	−0.004	0.027	−0.007	0.055
				β_2	3.0000	3.0070	0.700	0.498	0.233	0.166
				σ_1^2	0.0025	0.0025	−0.001	0.001	**−0.241**	0.442
				σ_2^2	0.5630	0.5625	−0.047	0.244	−0.083	0.434
				ρ	0.0000	−0.0007	−0.066	0.274	–	–
33	0.8	0.5	1000	β_1	0.5000	0.5008	0.082	0.064	0.164	0.128
				β_2	3.0000	2.9963	−0.375	0.849	−0.125	0.283
				σ_1^2	0.0025	0.0025	0.000	0.003	0.085	1.053
				σ_2^2	0.5630	0.5725	**0.946**	0.753	**1.680**	1.338
				ρ	0.5000	0.5041	0.414	0.598	0.828	1.197
34	0.8	0.5	5000	β_1	0.5000	0.5000	0.001	0.026	0.002	0.053
				β_2	3.0000	2.9979	−0.207	0.382	−0.069	0.127
				σ_1^2	0.0025	0.0025	−0.001	0.001	−0.557	0.545
				σ_2^2	0.5630	0.5586	−0.444	0.269	**−0.788**	0.477
				ρ	0.5000	0.4982	−0.178	0.289	−0.357	0.579
35	0.8	0.5	1000	β_1	0.5000	0.4992	**−0.083**	0.057	−0.166	0.114
				β_2	3.0000	2.9924	**−0.756**	0.892	−0.252	0.297
				σ_1^2	0.0025	0.0025	−0.001	0.002	−0.543	0.957
				σ_2^2	0.5630	0.5565	−0.652	0.372	**−1.158**	0.660
				ρ	0.8000	0.7965	−0.348	0.165	−0.435	0.206
36	0.8	0.5	5000	β_1	0.5000	0.5003	0.032	0.019	0.064	0.039
				β_2	3.0000	3.0071	0.705	0.416	0.235	0.139
				σ_1^2	0.0025	0.0025	−0.001	0.001	**−0.405**	0.400
				σ_2^2	0.5630	0.5610	−0.199	0.201	−0.354	0.357
				ρ	0.8000	0.7996	−0.036	0.106	−0.045	0.132

Table 3.7: Mean of estimates, estimated bias, relative bias and their estimated standard errors for different data size and correlations with $\beta_1 < \beta_2$ and $PSNR_1 > PSNR_2$

For the non-zero values of ρ we observe relative biases between -0.63% and -0.21%, whereas we have an estimated bias of -0.55% for $n = 1000$ and of -0.41% for $n = 5000$ when Y_{i1} and Y_{i2} are uncorrelated.

For scenarios 31 to 36, we exchange the values of $PSNR_1$ and $PSNR_2$. In those scenarios with $\beta_1 = \beta_2$, it was not necessary to consider this case. But here, we have to distinguish between the situation when the low $PSNR_j$ belongs to the low β_j, and when the low $PSNR_j$ belongs to the high β_j. The results for the latter case are provided in Table 3.7 and can also be compared to those in Table 3.3 where $\beta_1 = \beta_2$ holds.

We find the maximum absolute values of the estimated bias of both β_1 and β_2 in scenario 35 with high correlation and $n = 5000$. The corresponding values for σ_1^2 and σ_2^2 can be observed in the scenarios 31 and 33, where $n = 1000$ holds.

Looking at the maximum absolute value of the relative bias for each scenario in Table 3.7, we observe values ranging from 0.24% to 1.68%, which all correspond to the parameters σ_1^2 or σ_2^2.

Sc. #	$PSNR_1$	$PSNR_2$	n	θ	θ_{true}	$\bar{\theta}_{mod}$	$10^2 \cdot \widehat{b}(\widehat{\theta}_{mod})$	$10^2 \cdot s(\bar{\theta}_{mod})$	$10^2 \cdot \widehat{rb}(\widehat{\theta}_{mod})$	$10^2 \cdot s_{rb}(\widehat{\theta}_{mod})$
37	0.8	0.8	1000	β_1	0.5000	0.5011	**0.112**	0.072	0.223	0.144
				β_2	3.0000	3.0006	0.060	0.479	0.020	0.160
				σ_1^2	0.0025	0.0025	0.001	0.002	0.393	0.807
				σ_2^2	0.0900	0.0891	**−0.093**	0.097	**−1.029**	1.081
				ρ	0.0000	0.0064	0.640	0.666	–	–
38	0.8	0.8	5000	β_1	0.5000	0.5004	0.037	0.025	0.073	0.051
				β_2	3.0000	3.0019	0.192	0.137	0.064	0.046
				σ_1^2	0.0025	0.0025	−0.001	0.001	**−0.226**	0.438
				σ_2^2	0.0900	0.0900	0.004	0.041	0.048	0.457
				ρ	0.0000	−0.0032	−0.323	0.282	–	–
39	0.8	0.8	1000	β_1	0.5000	0.5007	0.072	0.056	0.144	0.113
				β_2	3.0000	3.0025	0.250	0.355	0.083	0.118
				σ_1^2	0.0025	0.0025	**−0.003**	0.002	−1.129	0.959
				σ_2^2	0.0900	0.0895	−0.049	0.073	−0.545	0.812
				ρ	0.5000	0.4854	−1.464	0.682	**−2.928**	1.364
40	0.8	0.8	5000	β_1	0.5000	0.4997	−0.027	0.033	−0.055	0.065
				β_2	3.0000	2.9990	−0.105	0.166	−0.035	0.055
				σ_1^2	0.0025	0.0025	−0.001	0.001	**−0.533**	0.355
				σ_2^2	0.0900	0.0900	−0.004	0.029	−0.047	0.328
				ρ	0.5000	0.4988	−0.118	0.268	−0.237	0.536
41	0.8	0.8	1000	β_1	0.5000	0.5005	0.050	0.054	0.100	0.107
				β_2	3.0000	3.0046	**0.458**	0.385	0.153	0.128
				σ_1^2	0.0025	0.0025	−0.000	0.003	−0.051	1.213
				σ_2^2	0.0900	0.0899	−0.014	0.099	**−0.154**	1.095
				ρ	0.8000	0.7997	−0.027	0.302	−0.034	0.377
42	0.8	0.8	5000	β_1	0.5000	0.4998	−0.017	0.026	−0.034	0.052
				β_2	3.0000	2.9990	−0.102	0.172	−0.034	0.057
				σ_1^2	0.0025	0.0025	−0.000	0.001	−0.119	0.428
				σ_2^2	0.0900	0.0900	−0.002	0.036	−0.017	0.403
				ρ	0.8000	0.7989	−0.110	0.116	**−0.138**	0.145

Table 3.8: Mean of estimates, estimated bias, relative bias and their estimated standard errors for different data size and correlations with $\beta_1 < \beta_2$ and high $PSNR_1 = PSNR_2$

For the correlation parameters $\rho \neq 0$, we get estimates of the relative bias from -0.44% to 0.83% and in the cases with $\rho = 0$, we get the values 0.25% and 0.07% for the estimated bias.

Scenarios with different β's and high $PSNR_1 = PSNR_2$

The last "block" of scenarios we look at are those when the β_j parameters are different and both $PSNR_j$ values are fixed to 80%. The results of the MCMC parameter estimation are provided in Table 3.8.

The maximum absolute value of the estimated bias of β_1 is 0.11% and located in scenario 37, that of β_2 is 0.46% and belongs to scenario 41. For the σ_j^2 parameters we find the corresponding maximums in scenarios 39 and 37. In all four cases, $n = 1000$ holds.

The estimated relative bias reaches absolute values up to 2.33%, which belongs to parameter ρ and scenario 39. The relative deviations of the ρ estimates in the other scenarios are small. For the independence scenarios 37 and 38, we have estimated biases of 0.64% and -0.32%. Over all parameters, the absolute values of the relative bias in Table 3.8 drop in general when the data size n is decreased.

Comparison of all scenarios

By looking at the results over all scenarios, one can say that the algorithm provides good estimates for all parameters in every considered situation. We can see this by looking at the relative bias. In most scenarios, the relative bias and its standard error have a scale of 10^{-2} or even 10^{-3}, and there are almost as much situations when the relative bias is positive as when it is negative. The maximum absolute value of the relative bias is only 2.93%, the second largest one is 1.75%. While this maximum can be found in scenario 39 where the data size is $n = 1000$, we can see that rising the sample size to $n = 5000$ leads to a much lower estimated relative bias (see scenario 40).

Overall effect of n

This is an effect we observe for most of the scenarios. There are a few scenarios where the absolute value of relative bias of one or more parameters is greater for $n = 5000$ than in the corresponding scenario with $n = 1000$, but this may be due to randomness and only 20 replications. The fact that in 73 out of 98 pairs the relative bias is greater for $n = 1000$ than for $n = 5000$ and also the higher standard error indicate consistence of the MCMC estimators. For most of the parameters, the relative bias decreases about $50 - 60\%$.

Overall effect of $PSNR_j$

Next, we look at the effect of a different signal-to-noise ratio, at first for the case when this ratio is the same for Y_{i1} and Y_{i2}. When $PSNR_j$ increases from 50% to 80%, i. e. when there are more data points where the signal dominates the noise, one can observe in the results that the absolute value of the relative bias gets smaller, i. e. the algorithm provides closer estimates. This especially holds for the $\boldsymbol{\beta}$ estimates, which is corresponding to what we expect, as it is easier to estimate the trend parameters when the signal dominates the noise. For the $\boldsymbol{\sigma}^2$ and ρ estimators this effect is less clear, but results suggest the possibility of an effect on the error behavior. Looking at the standard error of the $\boldsymbol{\beta}$-estimates we see that it clearly drops in every considered scenario when the $PSNR_j$ values are increased. This is a result of the lower values of $\boldsymbol{\sigma}^2$. For the other parameters this error stays quite similar when the signal-to-noise ratio changes. Averaging over all scenarios with $PSNR_1 = PSNR_2$ and all parameters, the absolute value of the relative bias drops from 0.38% to 0.29%.

In those scenarios with different $PSNR_j$ for Y_{i1} and Y_{i2}, we observe that the β_j-parameter with higher $PSNR_j$ often has a lower absolute value of the relative bias, whereas its standard error is always smaller than that of the other parameter β_k. Also in most of the cases the relative bias of the σ_j^2 with the higher $PSNR_j$ has a lower absolute value, but their standard errors hardly differ. This situation inside of a scenario is quite similar to the situation before, where we compared different scenarios to each other. It is also interesting to compare scenarios $25 - 30$ to scenarios $31 - 36$. In all of those scenarios it holds $PSNR_1 \neq PSNR_2$, but in the first group the lower β_j has the smaller $PSNR_j$, whereas it has the greater $PSNR_j$ in the second group. That means we compare for instance the relative bias of β_1 in scenario 25 to that of β_2 in scenario 31. We can see

that the relative biases change, but there is no clear direction: In some cases, the relative bias goes up, and in other cases it goes down. If we include the standard errors in our consideration, the the bias change gets less meaningful.

Overall effect of the correlation ρ

When the true value for the correlation parameter ρ is increased from 0 via 0.5 to 0.8, this has no systematic effect on the estimates of the relative bias of β_j. For σ_j^2, there also is no observable overall rise or fall when ρ is increased. As mentioned before, the standard error of the relative bias of β_j shows some dependance on the value of σ_j^2, but we see no big changes of it when ρ is increased. Furthermore, also the standard errors of the relative bias of σ_j^2 stay at least similar when only the correlation ρ changes. For the relative bias of ρ itself we see a clearer effect: In most comparable scenarios its absolute value drops when ρ is increased from 0.5 to 0.8.

Overall effect of the regression parameter $\boldsymbol{\beta}$

At last, we look at the impact of changing the true values of $\boldsymbol{\beta}^2$. As described before, for each scenario where $\beta_1 = \beta_2$ holds, we have another scenario where $\beta_1 < \beta_2$ holds but all other parameters stay unchanged. If we compare the relative biases, we see hardly any relationship between them. In some cases the relative bias rises, in others it drops or stays approximately the same. However, the standard errors of the relative bias roughly stay the same when β_2 is increased.

Chapter 4

Multivariate regression normal copula model with a single common covariate

4.1 Introduction

After we looked intensively at the bivariate case in the last chapter, we now want to develop and investigate a multivariate model, which allows us to consider for example d asset classes given the market return. The approach works similarly to that in the bivariate case, but here the main difficulty is handling the dependence.

4.2 Model definition

We look for a model for a dataset consisting of n response vectors $\boldsymbol{y}_i = (y_{i1}, \ldots, y_{id})'$ and n observations $z_1, \ldots, z_n$ $(i \in \{1, \ldots, n\})$. Similarly to the bivariate case, we choose a regression model with correlated, normal distributed errors.

Let $n \in \mathbb{N}$, $d \in \mathbb{N}$, $\boldsymbol{\beta} \in \mathbb{R}^d$, $\boldsymbol{\sigma}^2 \in (0, \infty)^d$ and let $R = (\rho_{jk})_{j,k=1,\ldots,d}$ be a correlation matrix. Assume that $\boldsymbol{\varepsilon}_i := (\varepsilon_{i1}, \ldots, \varepsilon_{id})' \sim \mathcal{N}_d(0, R)$ i.i.d. $\forall i = 1, \ldots, n$.

For given values $z_1, \ldots, z_d$ we define the multivariate model by

$$Y_{ij} = z_i \beta_j + \sigma_j \varepsilon_{ij} \quad (i = 1, \ldots n)\ (j = 1, \ldots, d) \tag{4.1}$$

In the multivariate case, it is helpful to work with vectors and matrices. With the definitions $\boldsymbol{Y}_i := (Y_{i1}, \ldots, Y_{id})'$, $X_i = z_i I_d$ (where I_d is the d-dimensional identity matrix) and $D := diag(\sigma_1, \ldots, \sigma_d)$, we can write (4.1) in the form

$$\boldsymbol{Y}_i = X_i \boldsymbol{\beta} + D\boldsymbol{\varepsilon}_i \tag{4.2}$$

4.3 Prior choices

Since we want to consider Bayesian estimates, we have to specify prior distributions for all our parameters. We choose these similarly to the bivariate case.
Assume all components $\beta_1, \ldots, \beta_d$ and $\sigma_1^2, \ldots, \sigma_d^2$ and the correlation matrix R are priorly independent and

$$\begin{aligned} \beta_j &\sim \mathcal{N}\left(0, s_j^2\right) && \text{independent} \quad \forall j = 1, \ldots, d \\ \sigma_j^2 &\sim InverseGamma(a_j, b_j) && \text{independent} \quad \forall j = 1, \ldots, d \\ p(R) &\propto 1 \text{ (which is a proper prior)} \end{aligned}$$

where $s_1^2, \ldots, s_d^2 > 0$, $a_1, \ldots, a_d > 0$ and $b_1, \ldots, b_d > 0$ are parameters that can be chosen subject to the prior information.

As mentioned in Section 3.3, the choice of independent priors has the advantage that a change on the prior distribution of one parameter does not affect the prior distribution of the other parameters. For an explanation of the prior choice for $\boldsymbol{\beta}$ and $\boldsymbol{\sigma}^2$, we refer to Section 3.3. The prior definition of the correlation matrix R is also consistent to the bivariate case, where we also chose a non-informative prior for the correlation parameter ρ. However, it is not clear in higher dimensions that we get a proper prior for the correlation matrix R. We see this later when we express R by a set of partial correlations.

The prior specification leads to the joint prior density

$$\begin{aligned} p(\boldsymbol{\beta}, \boldsymbol{\sigma}, \rho) &= \prod_{j=1}^{d} p(\beta_j) \cdot \prod_{j=1}^{d} p(\sigma_j^2) \cdot p(R) \\ p(\beta_j) &= \frac{1}{\sqrt{2\pi s_j^2}} \exp\left\{-\frac{\beta_j^2}{2s_j^2}\right\} && \forall j = 1, \ldots, d \\ p(\sigma_j^2) &= \frac{b_j^{a_j}}{\Gamma(a_j)} (\sigma_j^2)^{-a_j - 1} \exp\left\{-\frac{b_j}{\sigma_j^2}\right\} && \forall j = 1, \ldots, d \\ p(R) &\propto 1 \end{aligned}$$

4.4 Likelihood

Defining $\Sigma := DRD = \begin{pmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho_{12} & \cdots & \sigma_1\sigma_d\rho_{1d} \\ \sigma_1\sigma_2\rho_{12} & \sigma_2^2 & \cdots & \sigma_2\sigma_d\rho_{2d} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_1\sigma_d\rho_{1d} & \sigma_2\sigma_d\rho_{2d} & \cdots & \sigma_d^2 \end{pmatrix}$ we see from (4.2) that

$$\boldsymbol{Y}_i | \boldsymbol{\beta}, \Sigma \sim \mathcal{N}_d\left(X_i\boldsymbol{\beta}, \Sigma\right) \tag{4.3}$$

From the model definition (4.1) we know that $\boldsymbol{Y}_i$ is independent of $\boldsymbol{Y}_k$ for $i \neq k$. This means that the likelihood $f(\boldsymbol{y}|\boldsymbol{\beta}, \Sigma)$ is given by

$$f(\boldsymbol{y}|\boldsymbol{\beta}, \Sigma) = \prod_{i=1}^{n} \frac{1}{\sqrt{(2\pi)^d \det(\Sigma)}} \exp\left\{-\frac{1}{2}(\boldsymbol{y}_i - X_i\boldsymbol{\beta})'\Sigma^{-1}(\boldsymbol{y}_i - X_i\boldsymbol{\beta})\right\}$$

4.5 Full conditional distribution of $\boldsymbol{\beta}$

Now we want to determine the full conditional densities of all parameters, at first the full conditional density of $\boldsymbol{\beta}$.
From the prior specification in (4.1) we know that $\boldsymbol{\beta} \sim \mathcal{N}_d(0, \Gamma)$, where

$$\Gamma := \begin{pmatrix} s_1^2 & 0 & \cdots & 0 \\ 0 & s_2^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & s_d^2 \end{pmatrix}$$

Thus, the full conditional density for $\boldsymbol{\beta}$ is given by

$$\begin{aligned}
p(\boldsymbol{\beta}|D, R, \boldsymbol{y}) &\propto f(\boldsymbol{y}|\boldsymbol{\beta}, \Sigma)p(\boldsymbol{\beta}) \\
&\propto \left(\prod_{i=1}^{n} \exp\left\{-\frac{1}{2}(\boldsymbol{y}_i - X_i\boldsymbol{\beta})'\Sigma^{-1}(\boldsymbol{y}_i - X_i\boldsymbol{\beta})\right\}\right) \exp\left\{-\frac{1}{2}\boldsymbol{\beta}'\Gamma^{-1}\boldsymbol{\beta}\right\} \\
&\propto \exp\left\{ -\frac{1}{2}\left(\left(\sum_{i=1}^{n} 2\boldsymbol{y}_i'\Sigma^{-1}X_i\boldsymbol{\beta}\right) + \boldsymbol{\beta}'\left(\underbrace{\left(\sum_{i=1}^{n} X_i'\Sigma^{-1}X_i\right) + \Gamma^{-1}}_{=:\Gamma_n^{-1}} \right)\boldsymbol{\beta}\right)\right\} \\
&\propto \exp\left\{ -\frac{1}{2}\left(2\underbrace{\left(\sum_{i=1}^{n} \boldsymbol{y}_i'\Sigma^{-1}X_i\right)\Gamma_n}_{=:\boldsymbol{\mu}_n'} \Gamma_n^{-1}\boldsymbol{\beta} + \boldsymbol{\beta}'\Gamma_n^{-1}\boldsymbol{\beta}\right)\right\} \\
&\propto \exp\left\{-\frac{1}{2}(\boldsymbol{\beta} - \boldsymbol{\mu}_n)'\Gamma_n^{-1}(\boldsymbol{\beta} - \boldsymbol{\mu}_n)\right\}
\end{aligned} \tag{4.4}$$

Since $X_i = z_i I_d$, we can write for Γ_n^{-1}:

$$\Gamma_n^{-1} = \left(\sum_{i=1}^{n} X_i'\Sigma^{-1}X_i\right) + \Gamma^{-1} = \underbrace{\left(\sum_{i=1}^{n} z_i^2\right)}_{=:S_{zz}}\Sigma^{-1} + \Gamma^{-1} = S_{zz}\Sigma^{-1} + \Gamma^{-1}$$

So the expression for $\boldsymbol{\mu}_n$ can be transformed into

$$\boldsymbol{\mu}_n = \Gamma_n \left(\sum_{i=1}^{n} X_i \Sigma^{-1} \boldsymbol{y}_i \right) = \Gamma_n \left(\sum_{i=1}^{n} z_i \Sigma^{-1} \boldsymbol{y}_i \right) = \Gamma_n \Sigma^{-1} \Bigg(\underbrace{\sum_{i=1}^{n} z_i \boldsymbol{y}_i}_{=:S_{z\boldsymbol{y}}} \Bigg)$$

$$= \left(S_{zz} \Sigma^{-1} + \Gamma^{-1} \right)^{-1} \Sigma^{-1} S_{z\boldsymbol{y}} = \left(S_{zz} I_n + \Sigma \Gamma^{-1} \right)^{-1} S_{z\boldsymbol{y}}$$

Hence, the full conditional distribution of $\boldsymbol{\beta}$ is a normal distribution with mean $\boldsymbol{\mu}_n = (S_{zz} I_n + \Sigma \Gamma^{-1})^{-1} S_{z\boldsymbol{y}}$ and covariance matrix $\Gamma_n = (S_{zz} \Sigma^{-1} + \Gamma^{-1})^{-1}$.

Special case: If we choose $s_i^2 \to \infty$, i. e. $p(\boldsymbol{\beta}) \propto 1$, then $\Gamma^{-1} \to 0$, which leads to $\boldsymbol{\beta} | D, R, \boldsymbol{y} \sim \mathcal{N}_d \left(S_{zz}^{-1} S_{z\boldsymbol{y}}, S_{zz}^{-1} \Sigma \right)$

We are now able to construct an Update for $\boldsymbol{\beta}$ for the MCMC algorithm. This is done in Algorithm 4.1.

Algorithm 4.1 Update for $\boldsymbol{\beta}$

1: **INPUT** S_{zz}, $S_{z\boldsymbol{y}}$, $\boldsymbol{\sigma}^2$, Γ, R^{-1}
2: **OUTPUT** New value $\boldsymbol{\beta}_{new}$
3: $D := diag(\boldsymbol{\sigma})$ with $\boldsymbol{\sigma} := (\sigma_1, \ldots, \sigma_d)'$
4: $\Sigma^{-1} := D^{-1} R^{-1} D^{-1}$
5: $\Gamma_n := (S_{zz}^{-1} \Sigma^{-1} + \Gamma^{-1})^{-1}$
6: $\mu_n := \Gamma_n \Sigma^{-1} S_{z\boldsymbol{y}}$
7: Draw a sample $\boldsymbol{\beta}_{new}$ from $\mathcal{N}_d(\mu_n, \Gamma_n)$

4.6 Full conditional distribution of $\boldsymbol{\sigma}^2$

The next parameters of interest are the elements σ_j^2 of $\boldsymbol{\sigma}^2 = (\sigma_1^2, \ldots, \sigma_d^2)'$. Initially, we express the likelihood in terms of

$$V_{ij} = Y_{ij} - z_i \beta_j \quad \forall i = 1, \ldots n \; \forall j = 1, \ldots, d$$

We write $v_{ij} := y_{ij} - z_i \beta_j$, where y_{ij} are realizations of Y_{ij}.

From the properties of the multivariate normal distribution and the distribution of $\boldsymbol{Y}$ specified in (4.3) we know that $\boldsymbol{V}_i := (V_{i1}, \ldots, V_{id})' | \boldsymbol{\beta}, \Sigma \sim \mathcal{N}_d(0, \Sigma) \;\; \forall i = 1, \ldots, n$ and that they are independent, such that the likelihood can be expressed as

$$f(\boldsymbol{v} | \boldsymbol{\beta}, \Sigma) = \prod_{i=1}^{n} \frac{1}{\sqrt{(2\pi)^d \det(\Sigma)}} \exp \left\{ -\frac{1}{2} \boldsymbol{v}_i' \Sigma^{-1} \boldsymbol{v}_i \right\} \tag{4.5}$$

where $\boldsymbol{v}_i := (v_{i1}, \ldots, v_{id})'$ and $\boldsymbol{v} := (\boldsymbol{v}_1', \ldots, \boldsymbol{v}_n')'$.

We do not calculate the full conditional densities of $\sigma_1^2, \ldots, \sigma_d^2$ directly, but look instead at the inverse parametrization

$$\phi_j^2 := \frac{1}{\sigma_j^2} \quad \forall j = 1, \ldots, d$$

From the prior specification in (4.1) we know that for each $j = 1, \ldots, d$ it holds $\sigma_j^2 \sim InverseGamma(a_j, b_j)$, so it follows $\phi_j^2 \sim Gamma(a_j, b_j)$. We consider at first only the full conditional distribution of ϕ_1^2, the full conditional distributions of the parameters $\phi_2^2, \ldots, \phi_d^2$ are obtained similarly.

For that purpose we want to replace the covariance matrix $\Sigma = DRD$ in the likelihood (4.5) by $\phi_1^2, \ldots, \phi_d^2$ and the correlation matrix R. Let ρ^{jk} $(j, k = 1, \ldots, d)$ be the entries of the inverse correlation matrix R^{-1}.

We see that

$$\det(\Sigma) = \det(DRD) = \det(D)\det(R)\det(D) = \det(R)\prod_{j=1}^{d}\frac{1}{\phi_j^2} \tag{4.6}$$

and

$$\begin{aligned}\boldsymbol{v}_i'\Sigma^{-1}\boldsymbol{v}_i = \boldsymbol{v}_i'D^{-1}R^{-1}D^{-1}\boldsymbol{v}_i &= \sum_{j=1}^{d}\sum_{k=1}^{d}(D^{-1}\boldsymbol{v}_i)_j(D^{-1}\boldsymbol{v}_i)_k\rho^{jk} \\ &= \sum_{j=1}^{d}\sum_{k=1}^{d}\phi_j\phi_k\rho^{jk}v_{ij}v_{ik}\end{aligned} \tag{4.7}$$

By inserting (4.6) and (4.7) in the likelihood (4.5), we can derive the full conditional density of ϕ_1^2:

$$\begin{aligned}
&p(\phi_1^2|\phi_2^2, \ldots, \phi_d^2, \boldsymbol{\beta}, R, \boldsymbol{y}) \propto p(\phi_1^2)\prod_{i=1}^{n} f(\boldsymbol{v}_i|\boldsymbol{\beta}, \Sigma) \\
&= \frac{b_1^{a_1}}{\Gamma(a_1)}(\phi_1^2)^{a_1-1}\exp\{-b_1\phi_1^2\}\prod_{i=1}^{n}\frac{\sqrt{\prod_{j=1}^{d}\phi_j^2}}{\sqrt{(2\pi)^d\det(R)}}\exp\left\{-\frac{1}{2}\sum_{j=1}^{d}\sum_{k=1}^{d}\phi_j\phi_k\rho^{jk}v_{ij}v_{ik}\right\} \\
&\propto (\phi_1^2)^{a_1-1}\exp\{-b_1\phi_1^2\}\prod_{i=1}^{n}\sqrt{\phi_1^2}\exp\left\{-\frac{1}{2}\left(\phi_1^2\rho^{11}v_{i1}^2 + 2\phi_1\sum_{k=2}^{d}\phi_k\rho^{1k}v_{i1}v_{ik}\right)\right\} \\
&\propto (\phi_1^2)^{\frac{n-2}{2}+a_1}\exp\left\{-\frac{1}{2}\left(\phi_1^2(2b_1 + \rho^{11}\sum_{i=1}^{n}v_{i1}^2) + 2\sqrt{\phi_1^2}\sum_{k=2}^{d}\phi_k\rho^{1k}\sum_{i=1}^{n}v_{i1}v_{ik}\right)\right\} \\
&= (\phi_1^2)^{\frac{n-2}{2}+a_1}\exp\left\{-\frac{1}{2}\left(\phi_1^2(2b_1 + \rho^{11}S_{\boldsymbol{v},11}) + 2\sqrt{\phi_1^2}\sum_{k=2}^{d}\phi_k\rho^{1k}S_{\boldsymbol{v},1k}\right)\right\}
\end{aligned}$$

where $S_{\boldsymbol{v},jk} := \sum_{i=1}^{n} v_{ij} v_{ik}$ for $j, k = 1, \ldots, d$.

Performing similar calculations, one gets a proportional expression for the full conditional density of ϕ_j^2 for an arbitrary index $j \in \{1, \ldots, d\}$:

$$p(\phi_j^2 | \boldsymbol{\phi}_{-j}^2, \boldsymbol{\beta}, R, \boldsymbol{y}) \propto (\phi_j^2)^{\frac{n-2}{2}+a_j} \exp\left\{ -\frac{1}{2}\left(\phi_j^2 (2b_j + \rho^{jj} S_{\boldsymbol{v},jj}) + 2\sqrt{\phi_j^2} \sum_{\substack{k=1 \\ k\neq j}}^{d} \phi_k \rho^{jk} S_{\boldsymbol{v},jk} \right) \right\}$$
$$=: g_{\phi_j}(\phi_j^2, \boldsymbol{\phi}_{-j}^2, R^{-1}, S_{\boldsymbol{v}}) \tag{4.8}$$

where $\boldsymbol{\phi}_{-j}^2 := (\phi_1^2, \ldots, \phi_{j-1}^2, \phi_{j+1}^2, \ldots, \phi_d^2)'$ and $S_{\boldsymbol{v}} := (S_{\boldsymbol{v},jk})_{j,k=1,\ldots,d}$.

As in the bivariate case, we use the expression (4.8) to construct a Metropolis-Hastings step for $\boldsymbol{\sigma}^2$, by sequentially updating all components of $\boldsymbol{\phi}^2 := (\phi_1^2, \ldots, \phi_d^2)'$. As proposal density for each component ϕ_j^2, $j \in \{1, \ldots, d\}$, we choose again a normal distribution centered around the current value ϕ_j^2 with a predefined proposal variance s_{prop,ϕ_j}^2. The result is presented in Algorithm 4.2.
Due to numerical reasons, we also prefer using the logarithm of g_{ϕ_j} to calculate the acceptance probability. It holds

$$\log\left(g_{\phi_j}(\phi_j^2, \boldsymbol{\phi}_{-j}^2, R^{-1}, S_{\boldsymbol{v}})\right)$$
$$= \left(\frac{n-2}{2} + a_j\right) \log(\phi_j^2) - \frac{1}{2}\left(\phi_j^2 (2b_j + \rho^{jj} S_{\boldsymbol{v},jj}) + 2\sqrt{\phi_j^2} \sum_{\substack{k=1 \\ k\neq j}}^{d} \phi_k \rho^{jk} S_{\boldsymbol{v},jk} \right) \tag{4.9}$$

4.7 Full conditional distribution of the correlation matrix R

4.7.1 Reparametrization of R

At last, we want to determine the full conditional densities of all entries in the correlation matrix R. For that purpose, we use the "standardized" model residuals

$$\varepsilon_{ij} = \frac{Y_{ij} - z_i \beta_j}{\sigma_j} \quad \forall i = 1, \ldots n \; \forall j = 1, \ldots, d$$

From our model definition (4.1) we know that $\boldsymbol{\varepsilon}_i \sim \mathcal{N}_d(0, R) \;\; \forall i = 1, \ldots, n$ and that they are independent.

For $d \geq 3$ we have the problem that not every combination of $\{\rho_{jk} \in (-1, 1),\; 1 \leq j < k \leq d\}$ leads to admissible, i. e. positive definite values of R and Σ. To avoid this, we consider a d-dimensional D-vine $\mathcal{V}$ with constraint set $\mathcal{CV}$.

Algorithm 4.2 Update for $\boldsymbol{\sigma}^2$

INPUT $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$, $\boldsymbol{\sigma}^2$, $R^{-1} = (\rho^{jk})_{j,k=1,\ldots,d}$, $a_1, \ldots, a_d$, $b_1, \ldots, b_d$, $s^2_{prop,\phi_1}, \ldots, s^2_{prop,\phi_d}$
OUTPUT New value $\boldsymbol{\sigma}^2_{new}$
FOR $j := 1, \ldots, d$ **DO**
  FOR $k := j, \ldots, d$ **DO**
    $S_{\boldsymbol{v},jk} := S_{\boldsymbol{v},kj} := \sum_{i=1}^{n} v_{ij} v_{ik}$
  END FOR
END FOR
$S_{\boldsymbol{v}} := (S_{\boldsymbol{v},jk})_{j,k=1,\ldots,d}$
$\boldsymbol{\phi}^2 := \left(\frac{1}{\sigma_1^2}, \ldots, \frac{1}{\sigma_d^2}\right)'$
FOR $j := 1, \ldots, d$ **DO**
  Draw a sample $\phi^2_{prop,j}$ from $\mathcal{N}(\phi^2_j, s^2_{prop,\phi_j})$
  Calculate logarithm of acceptance probability
    $l_{acc} := \max\left\{\log(g_{\phi_j}(\phi^2_{prop,j}, \boldsymbol{\phi}^2_{-j}, R^{-1}, S_{\boldsymbol{v}})) - \log(g_{\phi_j}(\phi^2_j, \boldsymbol{\phi}^2_{-j}, R^{-1}, S_{\boldsymbol{v}})), 0\right\}$
    using formula (4.9)
  Draw a sample x from $Uniform(0, 1)$
  IF $x \leq \exp\{l_{acc}\}$ **THEN**
    $\phi^2_j := \phi^2_{prop,j}$
  END IF
END FOR
$\boldsymbol{\sigma}^2_{new} := \left(\frac{1}{\phi_1^2}, \ldots, \frac{1}{\phi_d^2}\right)'$

With the D-vine we can describe the correlation matrix R by a set of partial correlations

$$\{\rho_{jk;j+1:k-1} \in (-1, 1), \quad 1 \leq j < k \leq d\}$$

which is called a *complete partial correlation specification*. This means each partial correlation $\rho_{jk;j+1:k-1}$ corresponds to an edge $(jk|j+1:k-1)$ in the vine $\mathcal{V}$.
The advantage of this specification is that any arbitrary choice of $\rho_{jk;j+1:k-1} \in (-1, 1)$ leads to an admissible value of the correlation matrix R. For any d-dimensional regular vine $\mathcal{V}$ there exists a one-to-one correspondence between the set of d-dimensional correlation matrices and the set of partial correlation specifications for the vine (see Bedford and Cooke (2002), Corollary 7.5).

As we have seen in the preliminaries chapter, the likelihood of $\boldsymbol{\varepsilon}_i$ $(i = 1, \ldots, n)$ can now be formulated as follows:

$$\begin{aligned} &f(\varepsilon_{i1}, \ldots, \varepsilon_{id}|\boldsymbol{\rho}_{\mathcal{V}}) \\ &= \prod_{k=1}^{d} \varphi(\varepsilon_{ik}) \prod_{l=1}^{d-1} \prod_{j=1}^{d-l} c_{j,j+l|j+1:j+l-1}\left(F(\varepsilon_{ij}|\boldsymbol{\varepsilon}_{i,j+1:j+l-1}), F(\varepsilon_{i,j+l}|\boldsymbol{\varepsilon}_{i,j+1:j+l-1})|\rho_{j,j+l|j+1:j+l-1}\right) \end{aligned} \tag{4.10}$$

Next, we want to choose a prior density for $\boldsymbol{\rho}_{\mathcal{V}} := \{\rho_{jk;j+1:k-1},\ 1 \leq j < k \leq d\}$, such that the resulting distribution of R fits to our prior specification in (4.1).
For that purpose let $B_{(-1,1)}(\cdot, \alpha, \alpha)$ denote the density of a beta distributed random variable with equal parameters transformed from $(0,1)$ to $(-1,1)$, i. e.

$$B_{(-1,1)}(x, \alpha, \alpha) = \frac{2^{-2\alpha+1}\Gamma(2\alpha)}{(\Gamma(\alpha))^2}\left(1-x^2\right)^{\alpha-1} \tag{4.11}$$

Lewandowski, Kurowicka, and Joe (2007) propose the prior density

$$p\left(\boldsymbol{\rho}_{\mathcal{V}}\right) = \prod_{(jk|j+1:k-1)\in\mathcal{CV}} p\left(\rho_{jk;j+1:k-1}\right) \tag{4.12}$$

with

$$p\left(\rho_{jk;j+1:k-1}\right) = B_{(-1,1)}\left(\rho_{jk;j+1:k-1},\, \eta - 1 + \frac{d-(k-j-1)}{2},\, \eta - 1 + \frac{d-(k-j-1)}{2}\right) \tag{4.13}$$

where $\eta > 0$.

Furthermore, they show that the resulting prior density of R satisfies

$$p(R) \propto (\det(R))^{\eta-1}$$

By setting the parameter η to 1 in the prior specification (4.13), we get our non-informative prior for R defined in Section 4.3.

For our MCMC algorithm we can now sequentially update all components of $\boldsymbol{\rho}_{\mathcal{V}}$ by using MH-steps. For calculating the acceptance probability, we need an expression that is proportional to the likelihood of $\rho_{jk:j+1:k-1}$. To get such an expression, we have two possibilities.

4.7.2 Calculation with the D-vine likelihood

At first, we can calculate directly the D-vine likelihood of $\boldsymbol{\varepsilon} := (\boldsymbol{\varepsilon}_1', \ldots, \boldsymbol{\varepsilon}_n')'$ which depends on the copula parameters $\boldsymbol{\rho}_{\mathcal{V}}$. Since $\boldsymbol{\varepsilon}_1, \ldots, \boldsymbol{\varepsilon}_n$ are independent, it follows from (4.10):

$$\begin{aligned}
f(\boldsymbol{\varepsilon}|\boldsymbol{\rho}_{\mathcal{V}}) &= \prod_{i=1}^{n} f(\boldsymbol{\varepsilon}_i|\boldsymbol{\rho}_{\mathcal{V}}) \\
&= \prod_{i=1}^{n}\prod_{k=1}^{d}\varphi(\varepsilon_{ik}) \prod_{l=1}^{d-1}\prod_{j=1}^{d-l} c_{j,j+l|j+1:j+l-1}\left(F(\varepsilon_{ij}|\boldsymbol{\varepsilon}_{i,j+1:j+l-1}), F(\varepsilon_{i,j+l}|\boldsymbol{\varepsilon}_{i,j+1:j+l-1})|\rho_{j,j+l|j+1:j+l-1}\right) \\
&\propto \prod_{i=1}^{n}\prod_{l=1}^{d-1}\prod_{j=1}^{d-l} c_{j,j+l|j+1:j+l-1}\left(F(\varepsilon_{ij}|\boldsymbol{\varepsilon}_{i,j+1:j+l-1}), F(\varepsilon_{i,j+l}|\boldsymbol{\varepsilon}_{i,j+1:j+l-1})|\rho_{j,j+l|j+1:j+l-1}\right) \\
&=: g_{\boldsymbol{\rho}}(\boldsymbol{\rho}_{\mathcal{V}}, \boldsymbol{\varepsilon})
\end{aligned} \tag{4.14}$$

The advantage of this approach is that this expression depends only on all $u_{ij} := \Phi(\varepsilon_{ij})$ with $i \in \{1, \ldots, n\}$ and $j \in \{1, \ldots, d\}$: By definition, all copula densities do not depend on the marginal densities and as shown in Aas et al. (2009), also the conditional distribution functions can be calculated using only the copula parameters $\boldsymbol{\rho}_{\mathcal{V}}$ and the values of u_{ij}. So one can use any efficient algorithm which calculates D-vine likelihoods resp. log-likelihoods.

We get a proportional expression for the full conditional density of $\rho_{jk;j+1:k-1}$ by

$$\begin{aligned} p(\rho_{jk;j+1:k-1}|\rho_{jk;j+1:k-1}, \boldsymbol{\varepsilon}) &\propto f(\boldsymbol{\varepsilon}|\boldsymbol{\rho}_{\mathcal{V}}) p(\rho_{jk;j+1:k-1}) \\ &\propto g_{\boldsymbol{\rho}}(\boldsymbol{\rho}_{\mathcal{V}}, \boldsymbol{\varepsilon})(1 - \rho^2_{jk;j+1:k-1})^{\frac{d-k+j-1}{2}} \end{aligned}$$

The Metropolis-Hastings step for $\boldsymbol{\rho}_{\mathcal{V}}$ using the D-vine likelihood is provided in Algorithm 4.3. As proposal density for each element $\rho_{jk;j+1,k-1}$ of $\boldsymbol{\rho}_{\mathcal{V}}$ we use again a normal distribution with a predefined proposal variance $s^2_{prop,\rho_{jk;j+1:k-1}}$, centered around the current value of $\rho_{jk;j+1:k-1}$.

Algorithm 4.3 Update for $\boldsymbol{\rho}_{\mathcal{V}}$ using the D-vine likelihood

INPUT $\boldsymbol{\rho}_{\mathcal{V}}$, $\boldsymbol{\varepsilon}$, $\{s^2_{prop,\rho_{jk;j+1:k-1}}, 1 \leq j < k \leq d\}$
OUTPUT New value $\boldsymbol{\rho}_{new,\mathcal{V}}$
FOR $j := 1, \ldots, d$ **DO**
  FOR $k := j+1, \ldots, d$ **DO**
    Draw a sample ρ_{prop} from $\mathcal{N}(\rho_{jk;j+1:k-1}, s^2_{prop,\rho_{jk;j+1:k-1}})$
    Create a copy $\boldsymbol{\rho}_{prop,\mathcal{V}}$ of $\boldsymbol{\rho}_{\mathcal{V}}$, where the element with index $(jk; j+1:k-1)$ is replaced by $\rho_{prop,jk;j+1:k-1} := \rho_{prop}$
    Calculate logarithm of acceptance probability
      $l_{acc} := \log(g_{\boldsymbol{\rho}}(\boldsymbol{\rho}_{prop,\mathcal{V}}, \boldsymbol{\varepsilon})) - \log(g_{\boldsymbol{\rho}}(\boldsymbol{\rho}_{\mathcal{V}}, \boldsymbol{\varepsilon}))$
        $+ \frac{d-k+j-1}{2}\left(\log(1 - \rho^2_{prop,jk;j+1:k-1}) - \log(1 - \rho^2_{jk;j+1:k-1})\right)$
      using the definition of $g_{\boldsymbol{\rho}}$ in (4.14)
    Draw a sample x from $Uniform(0,1)$
    IF $x \leq \exp\{l_{acc}\}$ **THEN**
      $\boldsymbol{\rho}_{\mathcal{V}} := \boldsymbol{\rho}_{prop,\mathcal{V}}$
    END IF
  END FOR
END FOR
$\boldsymbol{\rho}_{new,\mathcal{V}} := \boldsymbol{\rho}_{\mathcal{V}}$

4.7.3 Calculation with the normal likelihood

Another possibility to get a proportional expression for the full conditional density of $\rho_{jk;j+1:k-1}$ is to calculate the correlation matrix R from the vector of partial correlations $\boldsymbol{\rho}_{\mathcal{V}}$. Recall that $\boldsymbol{\varepsilon}_i \sim \mathcal{N}_d(0, R)$, so

$$f(\boldsymbol{\varepsilon}_i|R) = \frac{1}{\sqrt{(2\pi)^d \det(R)}} \exp\left\{-\frac{1}{2}\boldsymbol{\varepsilon}_i' R^{-1} \boldsymbol{\varepsilon}_i\right\}$$

and the likelihood of $\boldsymbol{\varepsilon} = (\boldsymbol{\varepsilon}_1', \ldots, \boldsymbol{\varepsilon}_n')'$ can be written as

$$f(\boldsymbol{\varepsilon}|R) = \prod_{i=1}^{n} f(\boldsymbol{\varepsilon}_i|R) = (2\pi)^{-\frac{nd}{2}} \det(R)^{-\frac{n}{2}} \exp\left\{-\frac{1}{2}\sum_{i=1}^{n} \boldsymbol{\varepsilon}_i' R^{-1} \boldsymbol{\varepsilon}_i\right\} =: g_R(R, \boldsymbol{\varepsilon}) \quad (4.15)$$

With this normal likelihood and $\boldsymbol{\rho}_{\mathcal{V}\setminus(jk;j+1:k-1)}$ denoting the partial correlation vector $\boldsymbol{\rho}_{\mathcal{V}}$ without component $\rho_{jk;j+1:k-1}$, it holds

$$\begin{aligned} p(\rho_{jk;j+1:k-1}|\boldsymbol{\rho}_{\mathcal{V}\setminus(jk;j+1:k-1)}, \boldsymbol{\varepsilon}) &\propto f(\boldsymbol{\varepsilon}|R) p(\rho_{jk;j+1:k-1}) \\ &\propto g_R(R, \boldsymbol{\varepsilon})(1-\rho^2_{jk;j+1:k-1})^{\frac{d-k+j-1}{2}} \end{aligned}$$

where R is the correlation matrix derived from $\boldsymbol{\rho}_{\mathcal{V}}$ by using Algorithm 2.4. We summarize this alternative approach in Algorithm 4.4. In our implementation, we noticed that the update of Algorithm 4.4 is about 25% faster than the application of the D-vine likelihood in the update performed by Algorithm 4.3.

Algorithm 4.4 Update for $\boldsymbol{\rho}_{\mathcal{V}}$ using the normal likelihood

INPUT $\boldsymbol{\rho}_{\mathcal{V}}$, $\boldsymbol{\varepsilon}$, $\{s^2_{prop,\rho_{jk;j+1:k-1}}, 1 \le j < k \le d\}$
OUTPUT New value $\boldsymbol{\rho}_{new,\mathcal{V}}$
FOR $j := 1, \ldots, d$ **DO**
  FOR $k := j+1, \ldots, d$ **DO**
    Draw a sample ρ_{prop} from $\mathcal{N}(\rho_{jk;j+1:k-1}, s^2_{prop,\rho_{jk;j+1:k-1}})$
    Create a copy $\boldsymbol{\rho}_{prop,\mathcal{V}}$ of $\boldsymbol{\rho}_{\mathcal{V}}$, where the component $(jk; j+1:k-1)$ is replaced by $\rho_{prop,jk;j+1:k-1} := \rho_{prop}$
    Get correlation matrices R and R_{prop} from $\boldsymbol{\rho}_{\mathcal{V}}$ and $\boldsymbol{\rho}_{prop,\mathcal{V}}$ with Algorithm 2.4
    Calculate logarithm of acceptance probability
      $l_{acc} := \log(g_R(R_{prop}, \boldsymbol{\varepsilon})) - \log(g_R(R, \boldsymbol{\varepsilon}))$
        $+ \frac{d-k+j-1}{2}\left(\log(1-\rho^2_{prop,jk;j+1:k-1}) - \log(1-\rho^2_{jk;j+1:k-1})\right)$
      using the definition of g_R in (4.15)
    Draw a sample x from $Uniform(0,1)$
    IF $x \le \exp\{l_{acc}\}$ **THEN**
      $\boldsymbol{\rho}_{\mathcal{V}} := \boldsymbol{\rho}_{prop,\mathcal{V}}$
    END IF
  END FOR
END FOR
$\boldsymbol{\rho}_{new,\mathcal{V}} := \boldsymbol{\rho}_{\mathcal{V}}$

4.8 The multivariate MCMC algorithm

In the previous Sections 4.5 to 4.7 we developed updates for all parameters. With the derived algorithms, we are now able to construct an MCMC algorithm that samples from the joint posterior distribution. The result is – as in the bivariate case – a hybrid chain, since we have a Gibbs step for $\boldsymbol{\beta}$ and Metropolis Hastings steps for $\boldsymbol{\sigma}^2$ and for the partial correlations $\boldsymbol{\rho}_{\mathcal{V}}$ resp. the correlation matrix R. Algorithm 4.5 describes the procedure.

Algorithm 4.5 Multivariate MCMC algorithm

INPUT data $\boldsymbol{y}_i = (y_{i1}, \ldots, y_{id})'$ and z_i with $i = 1, \ldots, n$
  prior parameters $s_1^2, \ldots, s_d^2$, $a_1, \ldots, a_d$, $b_1, \ldots, b_d$
  number of MCMC iterations m (resp. size of posterior distribution sample)
  proposal variances $s^2_{prop,\phi_1}, \ldots, s^2_{prop,\phi_d}$ and $\{s^2_{prop,\rho_{jk;j+1:k-1}}, 1 \leq j < k \leq d\}$
  initial values $\boldsymbol{\beta}^{(0)}$, $\boldsymbol{\sigma}^{2(0)}$ and $\boldsymbol{\rho}_{\mathcal{V}}^{(0)}$
OUTPUT Samples $\boldsymbol{\beta}^{(1)}, \ldots, \boldsymbol{\beta}^{(m)}$, $\boldsymbol{\sigma}^{2(1)}, \ldots, \boldsymbol{\sigma}^{2(m)}$, $\boldsymbol{\rho}_{\mathcal{V}}^{(1)}, \ldots, \boldsymbol{\rho}_{\mathcal{V}}^{(m)}$

$\Gamma := diag(s_1, \ldots, s_d)$
$S_{zz} := \sum_{i=1}^n z_i^2$
$S_{z\boldsymbol{y}} := \sum_{i=1}^n z_i \boldsymbol{y}_i$
$\boldsymbol{\beta} := \boldsymbol{\beta}^{(0)}$, $\boldsymbol{\sigma}^2 := \boldsymbol{\sigma}^{2(0)}$, $\boldsymbol{\rho}_{\mathcal{V}} := \boldsymbol{\rho}_{\mathcal{V}}^{(0)}$
FOR $r := 1, \ldots, m$ **DO**
  Calculate correlation matrix R from $\boldsymbol{\rho}_{\mathcal{V}}$ using Algorithm 2.4
  Calculate inverse R^{-1}
  Calculate new value $\boldsymbol{\beta}^{(r)} := \boldsymbol{\beta}_{new}$ from Algorithm 4.1
  FOR $i := 1, \ldots, n$ **DO**
    $\boldsymbol{v}_i := \boldsymbol{y}_i - z_i \boldsymbol{\beta}^{(r)}$
  END FOR
  Calculate new value $\boldsymbol{\sigma}^{2(r)} := \boldsymbol{\sigma}^2_{new}$ from Algorithm 4.2
  $\boldsymbol{\sigma}^2 := \boldsymbol{\sigma}^{2(r)}$
  FOR $i := 1, \ldots, n$ **DO**
    FOR $j := 1, \ldots, d$ **DO**
      $\varepsilon_{ij} := \frac{v_{ij}}{\sigma_j}$
    END FOR
    $\boldsymbol{\varepsilon}_i := (\varepsilon_{i1}, \ldots, \varepsilon_{id})'$
  END FOR
  $\boldsymbol{\varepsilon} = (\boldsymbol{\varepsilon}_1', \ldots, \boldsymbol{\varepsilon}_n')'$
  Calculate new value $\boldsymbol{\rho}_{\mathcal{V}}^{(r)} := \boldsymbol{\rho}_{new,\mathcal{V}}$ from Algorithm 4.3
    or alternatively from Algorithm 4.4
  $\boldsymbol{\rho}_{\mathcal{V}} := \boldsymbol{\rho}_{\mathcal{V}}^{(r)}$
END FOR

4.9 The three dimensional case

In this section, we want to concentrate on $d = 3$ dimensions. The marginal parameters $\boldsymbol{\beta}$ and $\boldsymbol{\sigma}^2$ can be updated by using the formulas that we derived for the general case. However, the update of the correlation matrix takes a lot of time, so it makes sense to look for faster updates if one is only interested in the three dimensional case.
For that purpose, we consider the D-vine $\mathcal{V}$ with constraint set $\mathcal{CV} = \{(12), (23), (13|2)\}$. Recall that $\boldsymbol{\varepsilon}_i|\boldsymbol{\beta}, \Sigma \sim \mathcal{N}_3(0, R) \;\; \forall i = 1, \ldots, n$.

We can write the likelihood in terms of $\boldsymbol{\varepsilon}_i = (\varepsilon_{i1}, \varepsilon_{i2}, \varepsilon_{i3})'$ in the form

$$\begin{aligned} f(\boldsymbol{\varepsilon}_i|\boldsymbol{\beta}, \Sigma) = {} & c_{12}\left(F_1(\varepsilon_{i1}), F_2(\varepsilon_{i2})|\rho_{12}\right) c_{23}\left(F_2(\varepsilon_{i2}), F_3(\varepsilon_{i3})|\rho_{23}\right) \\ & \cdot c_{13|2}\left(F_{1|2}(\varepsilon_{i1}|\varepsilon_{i2}), F_{3|2}(\varepsilon_{i3}|\varepsilon_{i2})|\rho_{13;2}\right) f_1(\varepsilon_{i1}) f_2(\varepsilon_{i2}) f_3(\varepsilon_{i3}) \end{aligned} \tag{4.16}$$

where c_{12}, c_{23} and $c_{13|2}$ are Gauss copula densities with parameters ρ_{12}, ρ_{23} and $\rho_{13;2}$, which correspond to the respective (partial) correlations of $\boldsymbol{\varepsilon}_i$.
F_1, F_2 and F_3 denote the distribution functions of $\varepsilon_{i1}, \varepsilon_{i2}, \varepsilon_{i3}$ and f_1, f_2 and f_3 their density functions.
The conditional distribution function of ε_{i1} given ε_{i2} is $F_{1|2}$ and that of ε_{i3} given ε_{i2} is $F_{3|2}$.

Since $\varepsilon_{ij} \sim \mathcal{N}(0,1) \quad \forall i = 1, \ldots, n \; \forall j = 1, 2, 3$ we can write (4.16) as

$$\begin{aligned} f(\boldsymbol{\varepsilon}_i|\boldsymbol{\beta}, \Sigma) = {} & c_{12}\left(\Phi(\varepsilon_{i1}), \Phi(\varepsilon_{i2})|\rho_{12}\right) c_{23}\left(\Phi(\varepsilon_{i2}), \Phi(\varepsilon_{i3})|\rho_{23}\right) \\ & \cdot c_{13|2}\left(F_{1|2}(\varepsilon_{i1}|\varepsilon_{i2}), F_{3|2}(\varepsilon_{i3}|\varepsilon_{i2})|\rho_{13;2}\right) \varphi(\varepsilon_{i1})\varphi(\varepsilon_{i2})\varphi(\varepsilon_{i3}) \end{aligned} \tag{4.17}$$

where $\Phi(\cdot)$ denotes the standard normal distribution function and $\varphi(\cdot)$ the standard normal density function.

4.9.1 Full conditional distribution of ρ_{12}

Now we want to determine the full conditional density of ρ_{12}.
As a first step, we perform transformations of the likelihood $f(\boldsymbol{\varepsilon}_i, |\boldsymbol{\beta}, \Sigma)$ defined in (4.17), such that the resulting expression is proportional to $f(\boldsymbol{\varepsilon}_i|\boldsymbol{\beta}, \Sigma)$ in terms of ρ_{12}:

$$\begin{aligned} & f(\boldsymbol{\varepsilon}_i|\boldsymbol{\beta}, \Sigma) \\ & \quad \propto \underbrace{c_{12}\left(\Phi(\varepsilon_{i1}), \Phi(\varepsilon_{i2})|\rho_{12}\right) \varphi(\varepsilon_{i1})\varphi(\varepsilon_{i2})}_{=: g_{12}(\varepsilon_{i1}, \varepsilon_{i2}, \rho_{12})} c_{13|2}\left(F_{1|2}(\varepsilon_{i1}|\varepsilon_{i2}), F_{3|2}(\varepsilon_{i3}|\varepsilon_{i2})|\rho_{13;2}\right) \end{aligned} \tag{4.18}$$

Since $g_{12}(\varepsilon_{i1}, \varepsilon_{i2}, \rho_{12})$ has the form of a bivariate normal density with mean $(0,0)'$ and covariance matrix $\begin{pmatrix} 1 & \rho_{12} \\ \rho_{12} & 1 \end{pmatrix}$, it is equal to the joint density $f_{12}(\varepsilon_{i1}, \varepsilon_{i2})$ of ε_{i1} and ε_{i2}.

Thus we get

$$\begin{aligned} f\left(\boldsymbol{\varepsilon}_i|\boldsymbol{\beta},\Sigma\right) &\propto f_{12}\left(\varepsilon_{i1},\varepsilon_{i2}\right) c_{13|2}\left(F_{1|2}\left(\varepsilon_{i1}|\varepsilon_{i2}\right), F_{3|2}\left(\varepsilon_{i3}|\varepsilon_{i2}\right)|\rho_{13;2}\right) \\ &\propto c_{13|2}\left(F_{1|2}\left(\varepsilon_{i1}|\varepsilon_{i2}\right), F_{3|2}\left(\varepsilon_{i3}|\varepsilon_{i2}\right)|\rho_{13;2}\right) \underbrace{\frac{f_{12}\left(\varepsilon_{i1},\varepsilon_{i2}\right)}{f_2\left(\varepsilon_{i2}\right)}}_{=f_{1|2}(\varepsilon_{i1}|\varepsilon_{i2})} \underbrace{\frac{f_{23}\left(\varepsilon_{i2},\varepsilon_{i3}\right)}{f_2\left(\varepsilon_{i2}\right)}}_{=f_{3|2}(\varepsilon_{i3}|\varepsilon_{i2})} \\ &=: g_{13|2}\left(\boldsymbol{\varepsilon}_i,\rho_{12},\rho_{23},\rho_{13;2}\right) \end{aligned} \tag{4.19}$$

where $f_{23}\left(\varepsilon_{i2},\varepsilon_{i3}\right)$ denotes the joint density of $(\varepsilon_{i2},\varepsilon_{i3})'$ and $f_{1|2}(\varepsilon_{i1}|\varepsilon_{i2})$ resp. $f_{3|2}(\varepsilon_{i3}|\varepsilon_{i2})$ the conditional density of ε_{i1} resp. ε_{i3} given ε_{i2}.

Since

$$\begin{pmatrix}\varepsilon_{i1}\\ \varepsilon_{i2}\end{pmatrix} \sim \mathcal{N}_2\left(\begin{pmatrix}0\\0\end{pmatrix}, \begin{pmatrix}1 & \rho_{12}\\ \rho_{12} & 1\end{pmatrix}\right) \qquad \text{and} \qquad \begin{pmatrix}\varepsilon_{i2}\\ \varepsilon_{i3}\end{pmatrix} \sim \mathcal{N}_2\left(\begin{pmatrix}0\\0\end{pmatrix}, \begin{pmatrix}1 & \rho_{23}\\ \rho_{23} & 1\end{pmatrix}\right),$$

we know from (2.3) that $\varepsilon_{i1}|\varepsilon_{i2} \sim \mathcal{N}_1\left(\rho_{12}\varepsilon_{i2}, 1-\rho_{12}^2\right)$ and $\varepsilon_{i3}|\varepsilon_{i2} \sim \mathcal{N}_1\left(\rho_{23}\varepsilon_{i2}, 1-\rho_{23}^2\right)$, so $g_{13|2}\left(\boldsymbol{\varepsilon}_i,\rho_{12},\rho_{23},\rho_{13;2}\right)$ is a bivariate normal density with mean $\left(\rho_{12}\varepsilon_{i2},\rho_{23}\varepsilon_{i3}\right)'$ and covariance matrix

$$\begin{pmatrix} 1-\rho_{12}^2 & \rho_{13;2}\sqrt{1-\rho_{12}^2}\sqrt{1-\rho_{23}^2} \\ \rho_{13;2}\sqrt{1-\rho_{12}^2}\sqrt{1-\rho_{23}^2} & 1-\rho_{23}^2 \end{pmatrix}$$

This leads to

$$\begin{aligned} &f(\boldsymbol{\varepsilon}_i|\boldsymbol{\beta},\Sigma) \propto g_{13|2}\left(\boldsymbol{\varepsilon}_i,\rho_{12},\rho_{23},\rho_{13;2}\right) \\ &\overset{(2.1)}{=} \frac{1}{2\pi\sqrt{1-\rho_{12}^2}\sqrt{1-\rho_{23}^2}\sqrt{1-\rho_{13;2}^2}} \\ &\quad\cdot\exp\left\{-\frac{1}{2(1-\rho_{13;2}^2)}\left(\frac{(\varepsilon_{i1}-\rho_{12}\varepsilon_{i2})^2}{1-\rho_{12}^2}+\frac{(\varepsilon_{i3}-\rho_{23}\varepsilon_{i2})^2}{1-\rho_{23}^2}-\frac{2\rho_{13;2}(\varepsilon_{i1}-\rho_{12}\varepsilon_{i2})(\varepsilon_{i3}-\rho_{23}\varepsilon_{i2})}{\sqrt{1-\rho_{12}^2}\sqrt{1-\rho_{23}^2}}\right)\right\} \\ &\propto \frac{1}{\sqrt{1-\rho_{12}^2}}\exp\left\{-\frac{1}{2(1-\rho_{13;2}^2)\sqrt{1-\rho_{12}^2)}}\right. \\ &\qquad\left.\cdot\left(\frac{\varepsilon_{i1}^2-2\rho_{12}\varepsilon_{i1}\varepsilon_{i2}+\rho_{12}^2\varepsilon_{i2}^2}{\sqrt{1-\rho_{12}^2}}-\frac{2\rho_{13;2}(\varepsilon_{i1}\varepsilon_{i3}-\rho_{23}\varepsilon_{i1}\varepsilon_{i2}-\rho_{12}(\varepsilon_{i2}\varepsilon_{i3}-\rho_{23}\varepsilon_{i2}^2))}{\sqrt{1-\rho_{23}^2}}\right)\right\} \end{aligned} \tag{4.20}$$

By setting $j=1$, $k=2$ and $\eta=1$ in the prior definition (4.13) and by using (4.11), we get the prior density for ρ_{12}:

$$p(\rho_{12}) = \frac{2^{-2\cdot\frac{3}{2}+1}\Gamma(2\cdot\frac{3}{2})}{\left(\Gamma(\frac{3}{2})\right)^2}\left(1-\rho_{12}^2\right)^{\frac{3}{2}-1} \propto \left(1-\rho_{12}^2\right)^{\frac{1}{2}} \tag{4.21}$$

With (4.20) and the prior density (4.21) we can derive the full conditional density of ρ_{12}:

$$
\begin{aligned}
&p(\rho_{12}|\rho_{23},\rho_{13;2},\boldsymbol{\sigma}^2,\boldsymbol{\beta},\boldsymbol{y}) \propto p(\rho_{12})\left(\prod_{i=1}^{n} f\left(\boldsymbol{\varepsilon}_i,\boldsymbol{\beta},\Sigma\right)\right)\\
&\propto p(\rho_{12})\left(\prod_{i=1}^{n} g_{13|2}\left(\boldsymbol{\varepsilon}_i,\rho_{12},\rho_{23},\rho_{13;2}\right)\right)\\
&\propto \left(1-\rho_{12}^2\right)^{\frac{1}{2}}\prod_{i=1}^{n}\frac{1}{\sqrt{1-\rho_{12}^2}}\exp\left\{-\frac{1}{2(1-\rho_{13;2}^2)\sqrt{1-\rho_{12}^2}}\right.\\
&\qquad\left.\cdot\left(\frac{\varepsilon_{i1}^2-2\rho_{12}\varepsilon_{i1}\varepsilon_{i2}+\rho_{12}^2\varepsilon_{i2}^2}{\sqrt{1-\rho_{12}^2}}-\frac{2\rho_{13;2}(\varepsilon_{i1}\varepsilon_{i3}-\rho_{23}\varepsilon_{i1}\varepsilon_{i2}-\rho_{12}(\varepsilon_{i2}\varepsilon_{i3}-\rho_{23}\varepsilon_{i2}^2))}{\sqrt{1-\rho_{23}^2}}\right)\right\}\\
&=\left(1-\rho_{12}^2\right)^{-\frac{n-1}{2}}\exp\left\{-\frac{1}{2(1-\rho_{13;2}^2)\sqrt{1-\rho_{12}^2}}\left(\frac{\sum_{i=1}^{n}\varepsilon_{i1}^2-2\rho_{12}\sum_{i=1}^{n}\varepsilon_{i1}\varepsilon_{i2}+\rho_{12}^2\sum_{i=1}^{n}\varepsilon_{i2}^2}{\sqrt{1-\rho_{12}^2}}\right.\right.\\
&\qquad\left.\left.-\frac{2\rho_{13;2}(\sum_{i=1}^{n}\varepsilon_{i1}\varepsilon_{i3}-\rho_{23}\sum_{i=1}^{n}\varepsilon_{i1}\varepsilon_{i2}-\rho_{12}(\sum_{i=1}^{n}\varepsilon_{i2}\varepsilon_{i3}-\rho_{23}\sum_{i=1}^{n}\varepsilon_{i2}^2))}{\sqrt{1-\rho_{23}^2}}\right)\right\}
\end{aligned}
$$

By defining $S_{\boldsymbol{\varepsilon},jk} := \sum_{i=1}^{n}\varepsilon_{ij}\varepsilon_{ik}\ \forall j,k\in\{1,2,3\}$, we get

$$
\begin{aligned}
&p(\rho_{12}|\rho_{23},\rho_{13;2},\boldsymbol{\sigma}^2,\boldsymbol{\beta},\boldsymbol{y})\\
&\qquad\propto\left(1-\rho_{12}^2\right)^{-\frac{n-1}{2}}\exp\left\{-\frac{1}{2(1-\rho_{13;2}^2)\sqrt{1-\rho_{12}^2}}\left(\frac{S_{\boldsymbol{\varepsilon},11}-2\rho_{12}S_{\boldsymbol{\varepsilon},12}+\rho_{12}^2S_{\boldsymbol{\varepsilon},22}}{\sqrt{1-\rho_{12}^2}}\right.\right.\\
&\qquad\qquad\left.\left.-\frac{2\rho_{13;2}(S_{\boldsymbol{\varepsilon},13}-\rho_{23}S_{\boldsymbol{\varepsilon},12}-\rho_{12}(S_{\boldsymbol{\varepsilon},23}-\rho_{23}S_{\boldsymbol{\varepsilon},22}))}{\sqrt{1-\rho_{23}^2}}\right)\right\}
\end{aligned}
$$

Note that the expression above depends only on the parameters $\rho_{23},\rho_{13;2}$ and on the statistics $S_{\boldsymbol{\varepsilon},jk}$, which can be calculated fast.

4.9.2 Full conditional distribution of ρ_{23}

To obtain a proportional expression of the full conditional density of ρ_{23}, we can perform the same steps as in the previous derivation. This results in

$$
\begin{aligned}
&p(\rho_{23}|\rho_{12},\rho_{13;2},\boldsymbol{\sigma}^2,\boldsymbol{\beta},\boldsymbol{y})\\
&\qquad\propto\left(1-\rho_{23}^2\right)^{-\frac{n-1}{2}}\exp\left\{-\frac{1}{2(1-\rho_{13;2}^2)\sqrt{1-\rho_{23}^2}}\left(\frac{S_{\boldsymbol{\varepsilon},33}-2\rho_{23}S_{\boldsymbol{\varepsilon},23}+\rho_{23}^2S_{\boldsymbol{\varepsilon},22}}{\sqrt{1-\rho_{23}^2}}\right.\right.\\
&\qquad\qquad\left.\left.-\frac{2\rho_{13;2}(S_{\boldsymbol{\varepsilon},13}-\rho_{12}S_{\boldsymbol{\varepsilon},23}-\rho_{23}(S_{\boldsymbol{\varepsilon},12}-\rho_{12}S_{\boldsymbol{\varepsilon},22}))}{\sqrt{1-\rho_{12}^2}}\right)\right\}
\end{aligned}
$$

4.9.3 Full conditional distribution of $\rho_{13;2}$

At last, we determine the full conditional density of $\rho_{13;2}$. As before, we start with a transformation of the likelihood $f\left(\boldsymbol{\varepsilon}_i|\boldsymbol{\beta},\Sigma\right)$ defined in (4.17) that is proportional to the parameter of interest $\rho_{13;2}$:

$$\begin{aligned} f\left(\boldsymbol{\varepsilon}_i|\boldsymbol{\beta},\Sigma\right) &\propto c_{13|2}\left(F_{1|2}\left(\varepsilon_{i1}|\varepsilon_{i2}\right),F_{3|2}\left(\varepsilon_{i3}|\varepsilon_{i2}\right)|\rho_{13;2}\right) \\ &\propto c_{13|2}\left(F_{1|2}\left(\varepsilon_{i1}|\varepsilon_{i2}\right),F_{3|2}\left(\varepsilon_{i3}|\varepsilon_{i2}\right)|\rho_{13;2}\right)\underbrace{\frac{f_{12}(\varepsilon_{i1},\varepsilon_{i2})}{f_2(\varepsilon_{i2})}}_{=f_{1|2}(\varepsilon_{i1}|\varepsilon_{i2})}\underbrace{\frac{f_{23}(\varepsilon_{i2}|\varepsilon_{i3})}{f_2(\varepsilon_{i2})}}_{=f_{3|2}(\varepsilon_{i3}|\varepsilon_{i2})} \\ &= g_{13|2}\left(\boldsymbol{\varepsilon}_i,\rho_{12},\rho_{23},\rho_{13;2}\right) \end{aligned}$$

where we already know the expression $g_{13|2}\left(\boldsymbol{\varepsilon}_i,\rho_{12},\rho_{23},\rho_{13;2}\right)$ from (4.19). Note that $f_{12}(\varepsilon_{i1},\varepsilon_{i2})$ and $f_{23}(\varepsilon_{i2},\varepsilon_{i3})$ do depend on ρ_{12} and ρ_{23}, but not on $\rho_{13;2}$.

From the prior specification (4.13) with $j=1$, $k=3$ and $\eta=1$ and from (4.11) we receive

$$p(\rho_{13;2}) = \frac{2^{-2\cdot 1+1}\Gamma(2\cdot 1)}{\left(\Gamma(1)\right)^2}\left(1-\rho_{13;2}^2\right)^{1-1} \propto 1$$

so we have an a non-informative prior for $\rho_{13;2}$.

Thus, it follows that

$$\begin{aligned} p(\rho_{13;2}|\rho_{12},\rho_{23},\boldsymbol{\sigma}^2,\boldsymbol{\beta},\boldsymbol{y}) &\propto p(\rho_{13;2})\left(\prod_{i=1}^n f\left(\boldsymbol{\varepsilon}_i,\boldsymbol{\beta},\Sigma\right)\right) \propto \prod_{i=1}^n g_{13|2}\left(\boldsymbol{\varepsilon}_i,\rho_{12},\rho_{23},\rho_{13;2}\right) \\ &\propto \prod_{i=1}^n \frac{1}{\sqrt{1-\rho_{13;2}^2}}\exp\left\{-\frac{1}{2(1-\rho_{13;2}^2)}\left(\frac{(\varepsilon_{i1}-\rho_{12}\varepsilon_{i2})^2}{1-\rho_{12}^2}+\frac{(\varepsilon_{i3}-\rho_{23}\varepsilon_{i2})^2}{1-\rho_{23}^2}\right.\right. \\ &\qquad\left.\left.-\frac{2\rho_{13;2}(\varepsilon_{i1}-\rho_{12}\varepsilon_{i2})(\varepsilon_{i3}-\rho_{23}\varepsilon_{i2})}{\sqrt{1-\rho_{12}^2}\sqrt{1-\rho_{23}^2}}\right)\right\} \\ &\propto (1-\rho_{13;2}^2)^{-\frac{n}{2}}\exp\left\{-\frac{1}{2(1-\rho_{13;2}^2)}\left(\frac{\sum_{i=1}^n\varepsilon_{i1}^2-2\rho_{12}\sum_{i=1}^n\varepsilon_{i1}\varepsilon_{i2}+\rho_{12}^2\sum_{i=1}^n\varepsilon_{i2}^2}{1-\rho_{12}^2}\right.\right. \\ &\qquad+\frac{\sum_{i=1}^n\varepsilon_{i3}^2-2\rho_{23}\sum_{i=1}^n\varepsilon_{i2}\varepsilon_{i3}+\rho_{23}^2\sum_{i=1}^n\varepsilon_{i2}^2}{1-\rho_{23}^2} \\ &\qquad\left.\left.-\frac{2\rho_{13;2}(\sum_{i=1}^n\varepsilon_{i1}\varepsilon_{i3}-\rho_{23}\sum_{i=1}^n\varepsilon_{i1}\varepsilon_{i2}-\rho_{12}(\sum_{i=1}^n\varepsilon_{i2}\varepsilon_{i3}-\rho_{23}\sum_{i=1}^n\varepsilon_{i2}^2))}{\sqrt{1-\rho_{12}^2}\sqrt{1-\rho_{23}^2}}\right)\right\} \end{aligned}$$

Finally, we use again the notations $S_{\boldsymbol{\varepsilon},jk} = \sum_{i=1}^{n} \varepsilon_{ij}\varepsilon_{ik}\ \forall j,k \in \{1,2,3\}$ and get

$$
\begin{aligned}
&p(\rho_{13;2}|\rho_{12},\rho_{23},\boldsymbol{\sigma}^2,\boldsymbol{\beta},\boldsymbol{y}) \\
&\propto (1-\rho_{13;2}^2)^{-\frac{n}{2}} \exp\left\{-\frac{1}{2(1-\rho_{13;2}^2)}\left(\frac{S_{\boldsymbol{\varepsilon},11}-2\rho_{12}S_{\boldsymbol{\varepsilon},12}+\rho_{12}^2 S_{\boldsymbol{\varepsilon},22}}{1-\rho_{12}^2}\right.\right. \\
&\qquad \left.\left.+\frac{S_{\boldsymbol{\varepsilon},33}-2\rho_{23}S_{\boldsymbol{\varepsilon},23}+\rho_{23}^2 S_{\boldsymbol{\varepsilon},22}}{1-\rho_{23}^2}-\frac{2\rho_{13;2}(S_{\boldsymbol{\varepsilon},13}-\rho_{12}S_{\boldsymbol{\varepsilon},23}-\rho_{23}S_{\boldsymbol{\varepsilon},12}+\rho_{12}\rho_{23}S_{\boldsymbol{\varepsilon},22})}{\sqrt{1-\rho_{12}^2}\sqrt{1-\rho_{23}^2}}\right)\right\}
\end{aligned}
$$

4.10 Small sample performance using the multivariate MCMC algorithm for Bayesian inference

The algorithm was tested for different true parameter values with simulated data sets of different data sizes. For three dimensions, the results are provided in the appendix. The algorithm was also checked on higher dimensions, especially for dimension 12, since this dimension is relevant for our application in chapter 5.

The results were always acceptable and comparable to the bivariate case. However, we omit an extensive analysis of results here and refer to Section 3.7.

Chapter 5

Application: U.S. industrial returns

5.1 Data description

In this chapter we want to apply the MCMC algorithm for the multivariate regression normal copula model to real data. Our data consists of monthly excess returns of 12 different U.S. industrial portfolios such as manufacturing, energy or healthcare, and of the market return from July 1926 to December 2007, making a total of 978 observations. The returns were constructed by Kenneth R. French and sourced from the Wharton Data Service. With excess returns we mean that the risk-free rate is subtracted from each portfolio return and the market return.

To identify each portfolio, we use different symbols. All considered industrial returns and the corresponding portfolio symbols are listed in Table 5.1.

In accordance to our notation in the previous chapter, we denote the excess return of portfolio $j \in \{1, \ldots, 12\}$ at time $i \in \{1, \ldots 978\}$ by y_{ij}, and the market return at the same time as z_i. However, as we later change the order of the portfolios, we prefer to use the symbols N, D, etc. instead of the indices $1, \ldots, 12$. For example, y_{iN} denotes the return of the Consumer Non-Durables industry at time $i \in \{1, \ldots, 978\}$. As scale, we use per cent values.

At first, we look at the summary statistics of the monthly industrial returns and the market return given in per cent values in Table 5.2. One sees that most industry returns are in a range of -30 to 40 and that their means are positive. The quartiles also suggest a non symmetric distribution with positive skewness.

The impression of a skewed distribution for most portfolio returns is confirmed by the estimates of the centralized moments provided in Table 5.3. Besides the positive skewness we also see that estimated variance and estimated excess kurtosis differ clearly from each other. A high variance may result from a high leverage of the market return (i.e. high β_j) or a high deviation from the leveraged market return (i. e. high σ_j^2). With our model, we will be able to distinguish between these effects and additionally include correlation between residuals.

Symbol	Description / Type of industry
N	Consumer Non-Durables
D	Consumer Durables
M	Manufactoring
E	Energy
C	Chemicals
B	Business Equipment
T	Telecommunications
U	Utilities
S	Shops
H	Healthcare
\$	Money
O	Other

Table 5.1: Investigated industrial portfolios and their abbreviation

	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
z_i	-29.04	-0.21	0.97	0.65	3.66	38.27
y_{iN}	-24.51	-0.16	0.80	0.69	3.40	34.31
y_{iD}	-34.81	-0.30	0.70	0.80	4.48	79.61
y_{iM}	-29.15	-0.24	1.19	0.75	4.06	60.14
y_{iE}	-26.01	-0.25	0.65	0.80	4.25	33.47
y_{iC}	-31.41	-0.23	0.78	0.76	4.04	48.88
y_{iB}	-34.64	-0.32	0.90	0.82	4.86	58.47
y_{iT}	-21.59	-0.16	0.68	0.55	2.76	28.13
y_{iU}	-32.99	-0.20	0.76	0.61	3.27	43.13
y_{iS}	-30.21	-0.25	0.73	0.68	3.76	36.95
y_{iH}	-34.80	-0.22	0.74	0.79	3.82	38.56
$y_{i\$}$	-39.50	-0.23	0.85	0.77	4.15	59.82
y_{iO}	-31.19	-0.27	0.73	0.56	4.06	58.41

Table 5.2: Summary statistics of portfolio excess returns

	est. variance	est. skewness	est. kurtosis
z_i	29.3	0.22	7.95
y_{iN}	22.0	-0.01	5.90
y_{iD}	57.9	1.25	15.48
y_{iM}	45.6	1.12	13.35
y_{iE}	35.7	0.26	3.24
y_{iC}	34.4	0.49	8.88
y_{iB}	59.6	0.47	7.19
y_{iT}	20.9	0.07	3.37
y_{iU}	32.4	0.14	7.67
y_{iS}	34.8	-0.00	5.49
y_{iH}	33.4	0.18	7.18
$y_{i\$}$	46.8	0.66	12.31
y_{iO}	44.7	0.95	12.88

Table 5.3: Estimates of variance, skewness and (excess) kurtosis for the excess market return and the industrial portfolio returns

To get a further impression of the distribution of the excess market return, we look at its time series plot (Figure 5.1). We see that there is no clear trend, and that the volatility varies over time. Relatively large deviations from the mean can be observed around index 40 which corresponds to the Great Depression commencing in 1929 and the following economic recovery, around the beginning of the Second World War and at the time of the "Black Monday" in October 1987.

However, we are not interested in the distribution of the market return itself, but on its effect on the industrial portfolio returns. The strong dependence between the market return and the industrial returns can be observed by looking at the time series plots of all portfolio returns (see Figure 5.2). This apparently strong dependence suggests an application of the capital asset pricing model, i.e. a multivariate regression model with one common covariate.

Before we run the MCMC algorithm, we use the marginal estimators

$$\beta_j^{(mar)} := \frac{\sum_{i=1}^{n} z_i y_{ij}}{\sum_{i=1}^{n} z_i^2} \quad \text{and} \quad \sigma_j^{2(mar)} := \frac{\sum_{i=1}^{n} (y_{ij} - z_i \beta_j^{(mar)})^2}{n-1} \quad (j = 1, \ldots, 12)$$

to get residuals of the form

$$\varepsilon_{ij}^{(mar)} := \frac{y_{ij} - z_i \beta_j^{(mar)}}{\sigma_j^{2(mar)}} \quad (j = 1, \ldots, 12)$$

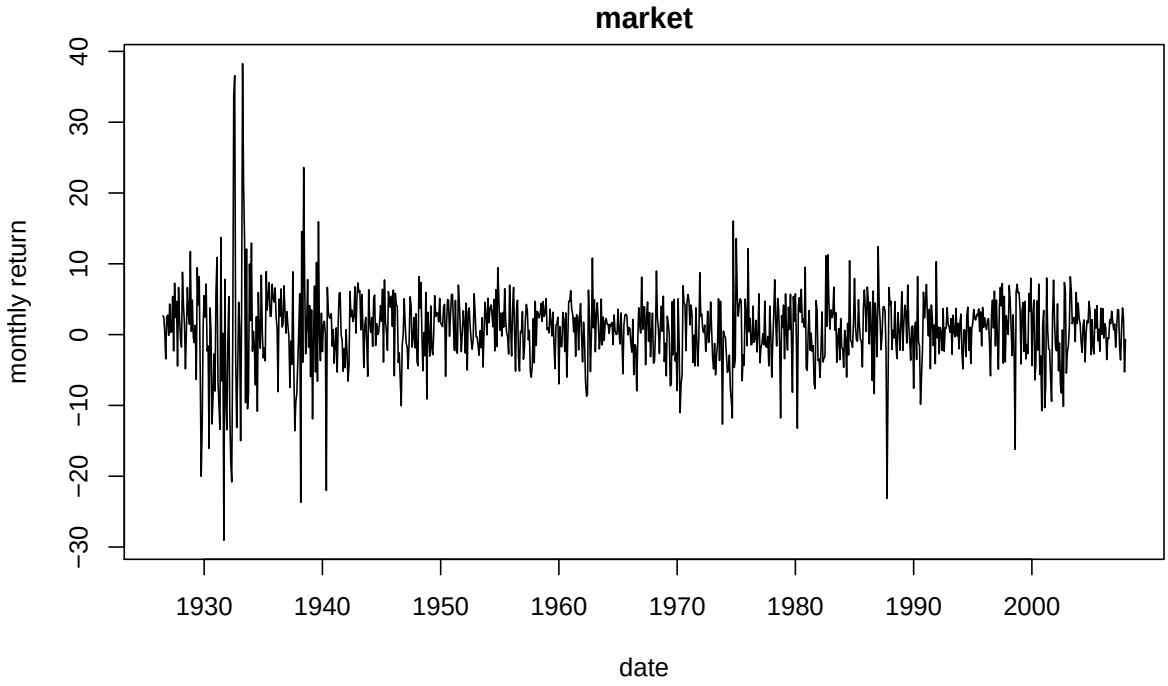

Figure 5.1: Time series plot of the excess market return

Using these residuals and the marginal estimates, we are able to calculate further quantities in the following Table 5.4 to learn more about the dataset. For each Portfolio index $j \in \{1, \ldots, 12\}$, we get besides the empirical standard deviation

$$\sigma_j^{(mar)} := \sqrt{\sigma_j^{2(mar)}}$$

an estimate of $PSNR_j$, which tells us how much of the data given the marginal estimates for β_j and σ_j^2 has a signal-to-noise ratio larger than 2.

Finally, we calculated for each portfolio $j \in \{1, \ldots, 12\}$ the goodness of fit measure $\widetilde{R}_j^2$, which is defined by

$$\widetilde{R}_j^2 := \frac{\sum_{i=1}^{n} (z_i \beta_j)^2}{\sum_{i=1}^{n} y_{ij}^2} \tag{5.1}$$

Note that this definition differs from the frequently used coefficient of determination R^2, since no intercept term is included in our model.

The marginal estimates are provided in Table 5.4. As we can see, the estimates for the regression parameters β_j and the residual variances σ_j^2 clearly differ among the portfolios. However, the values in the $PSNR_j$ column show that the signal-to-noise ratio suggested by the marginal estimation is often smaller 2, which means that the noise dominates the

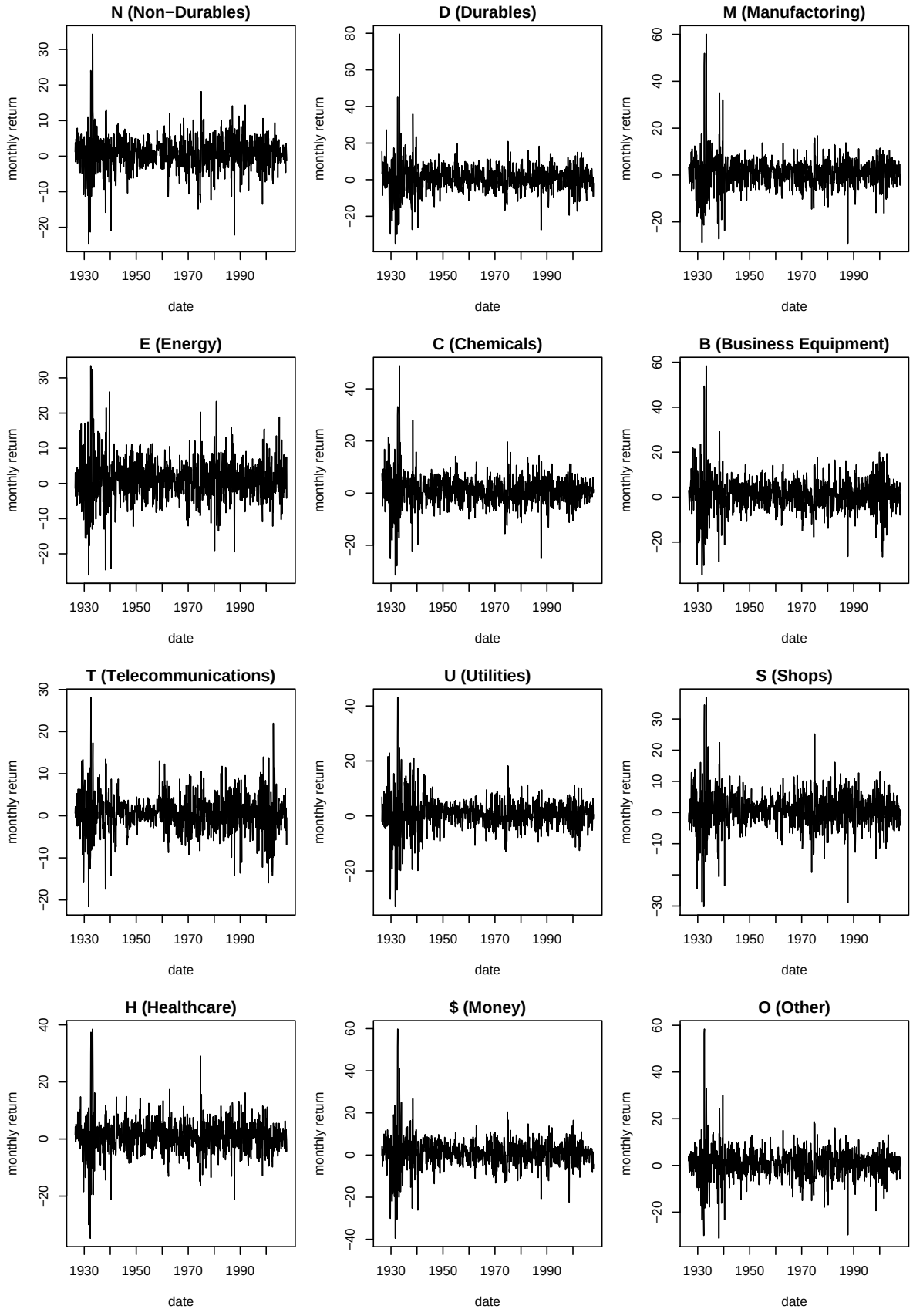

Figure 5.2: Time series plots of the industrial portfolios excess returns

Portfolio	$\beta_j^{(mar)}$	$\sigma_j^{2(mar)}$	$\sigma_j^{(mar)}$	$PSNR_j^{(mar)}$ in %	$\widetilde{R}_j^2$
N	0.77	5.02	2.24	19.02	0.78
D	1.22	14.43	3.80	16.05	0.75
M	1.19	4.08	2.02	44.58	0.91
E	0.86	14.41	3.80	6.85	0.60
C	0.98	6.55	2.56	23.21	0.81
B	1.29	10.52	3.24	25.26	0.83
T	0.64	8.94	2.99	6.34	0.58
U	0.81	13.46	3.67	6.34	0.59
S	0.96	7.83	2.80	19.22	0.78
H	0.87	11.80	3.44	9.00	0.65
\$	1.16	7.53	2.74	28.63	0.84
O	1.13	6.79	2.61	30.47	0.85

Table 5.4: Marginal estimates of β_j, σ_j^2, σ_j and $PSNR_j$ as well as the goodness of fit measure $\widetilde{R}_j^2$ for each portfolio

signal, which could lead to imprecise β_j estimates. When estimating those parameters with our MCMC approach, we can take a closer look on the accuracy of the estimates. In contrast to the low $PNSR_j$ values, the high $\widetilde{R}_j^2$ values of 60 to 91 per cent of the marginal regressions suggest that a high fraction of the total variance of y_{ij} is explained by the values of β_j and z_j.

The next subject of interest is the distribution of the standardized residuals, which we assume to be normal in our model. We take the standardized residuals $\varepsilon_{ij}^{(mar)}$ resulting from the marginal estimates and create QQ-plots shown in Figure 5.1 to find out more about their distribution. These plots show that the distribution of the residuals has heavier tails than the normal distribution, especially for portfolios M, C, \$ and O. However, the normal model is suitable to estimate the regression parameters, residual variances and the residual correlations. An alternative approach to our model is to assume a Student's t-distribution as marginal distribution for the standardized residuals ε_{ij}, as it is done in Pitt et al. (2006). We consider our model as a first model.

Another interesting point to analyze is if there is any autocorrelation in the residuals. In Figure 5.4 we therefore show plots of the empirical auto correlation function for the residuals of each portfolio and until a lag of 30. These plots propose that there is no correlation between portfolio returns at different points in time. This is what we expect, as it is the result of the unpredictability of excess returns: If there was a high autocorrelation between the residuals of the portfolio excess return, one could use the current return to predict whether the portfolio will outperform the market return in the next period.

The picture changes if one looks at the squares of the estimated residuals and their autocorrelation plots, as provided in Figure 5.5. One can see significant autocorrelation for almost all portfolio returns and for many lags.

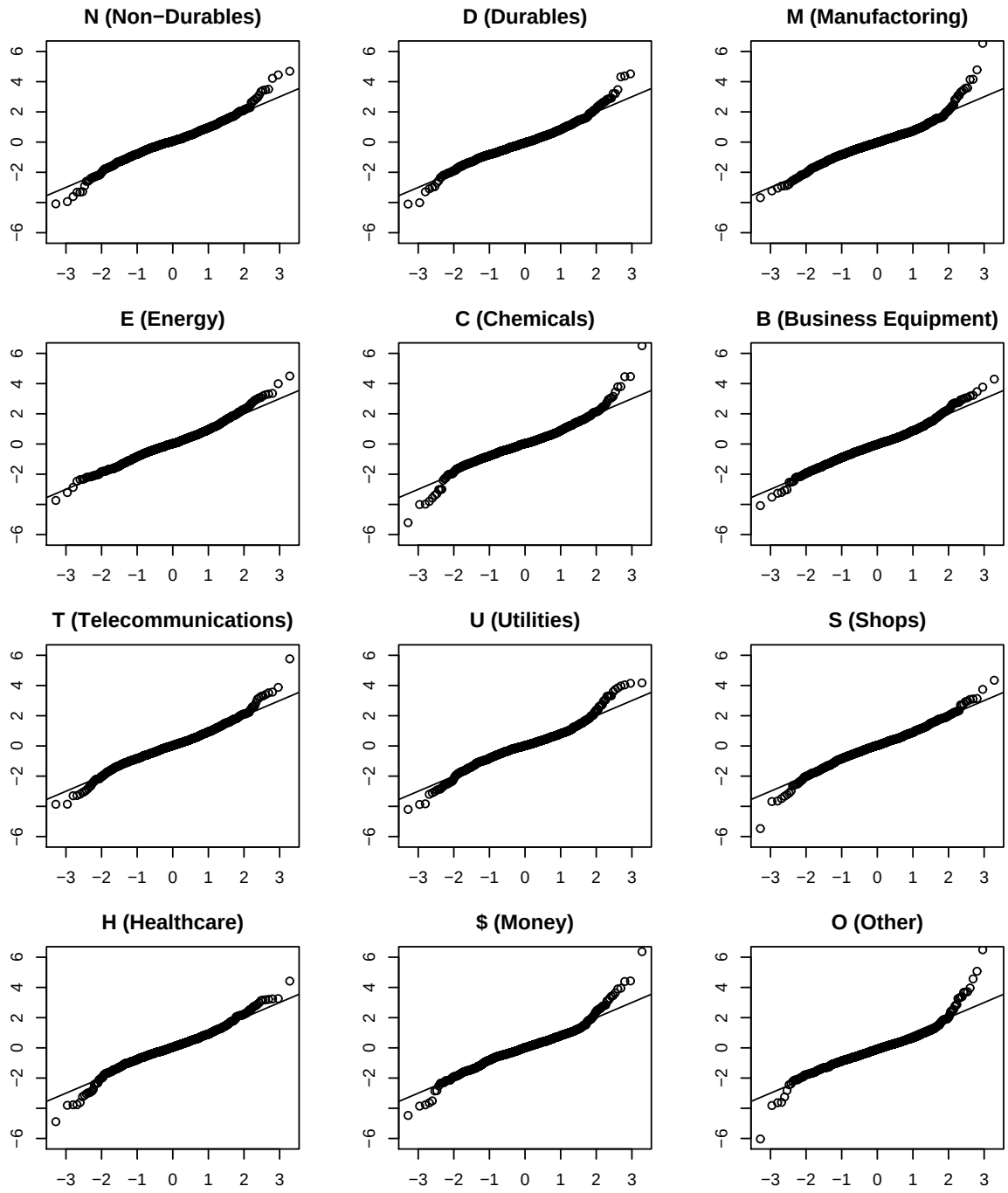

Figure 5.3: QQ plots of the standardized residuals $\varepsilon_{ij}^{(mar)}$ resulting from the marginal estimators $\boldsymbol{\beta}^{(mar)}$ and $\boldsymbol{\sigma}^{2(mar)}$

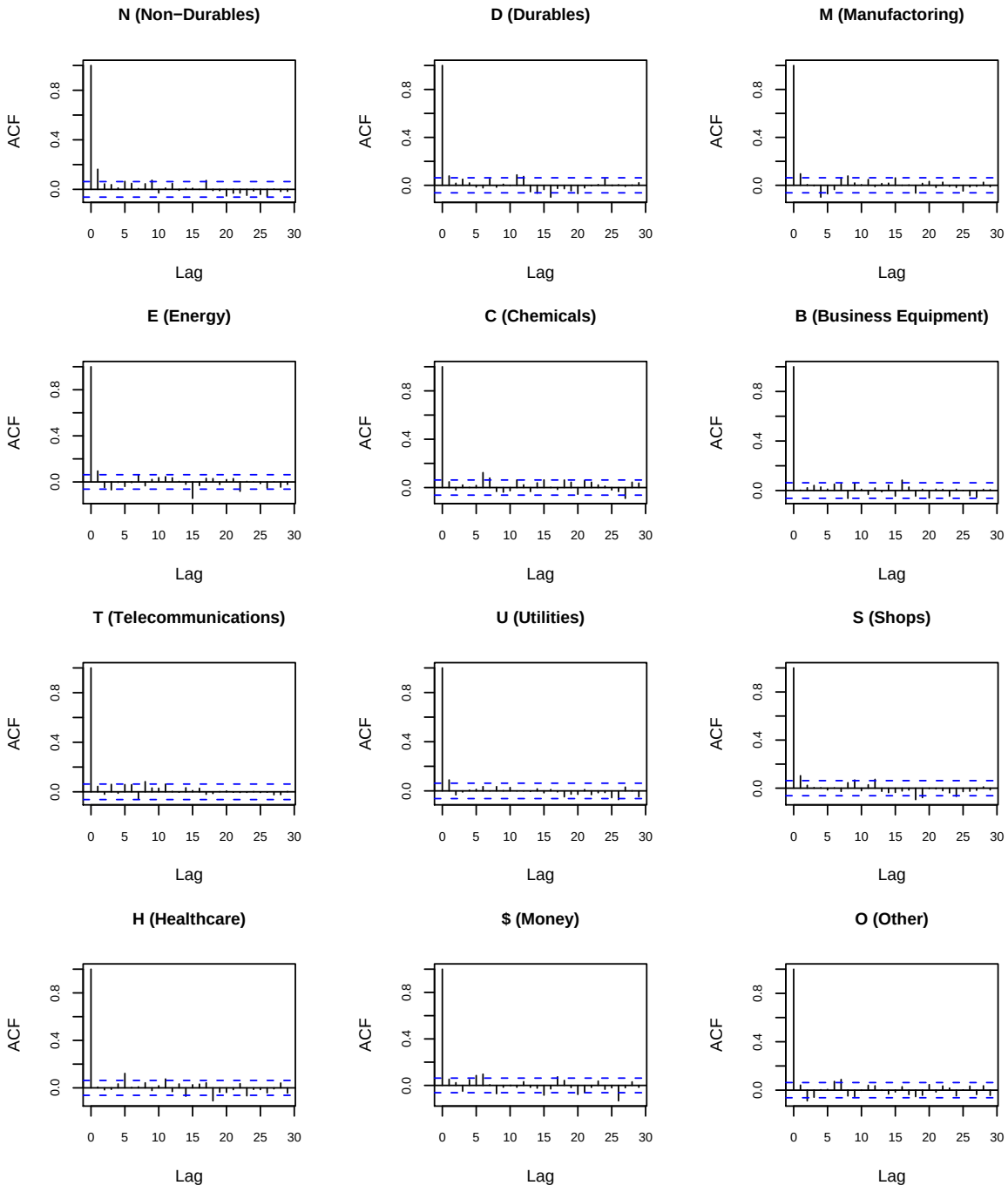

Figure 5.4: ACF plots of the standardized residuals $\varepsilon_{ij}^{(mar)}$ resulting from the marginal estimators $\boldsymbol{\beta}^{(mar)}$ and $\boldsymbol{\sigma}^{2(mar)}$

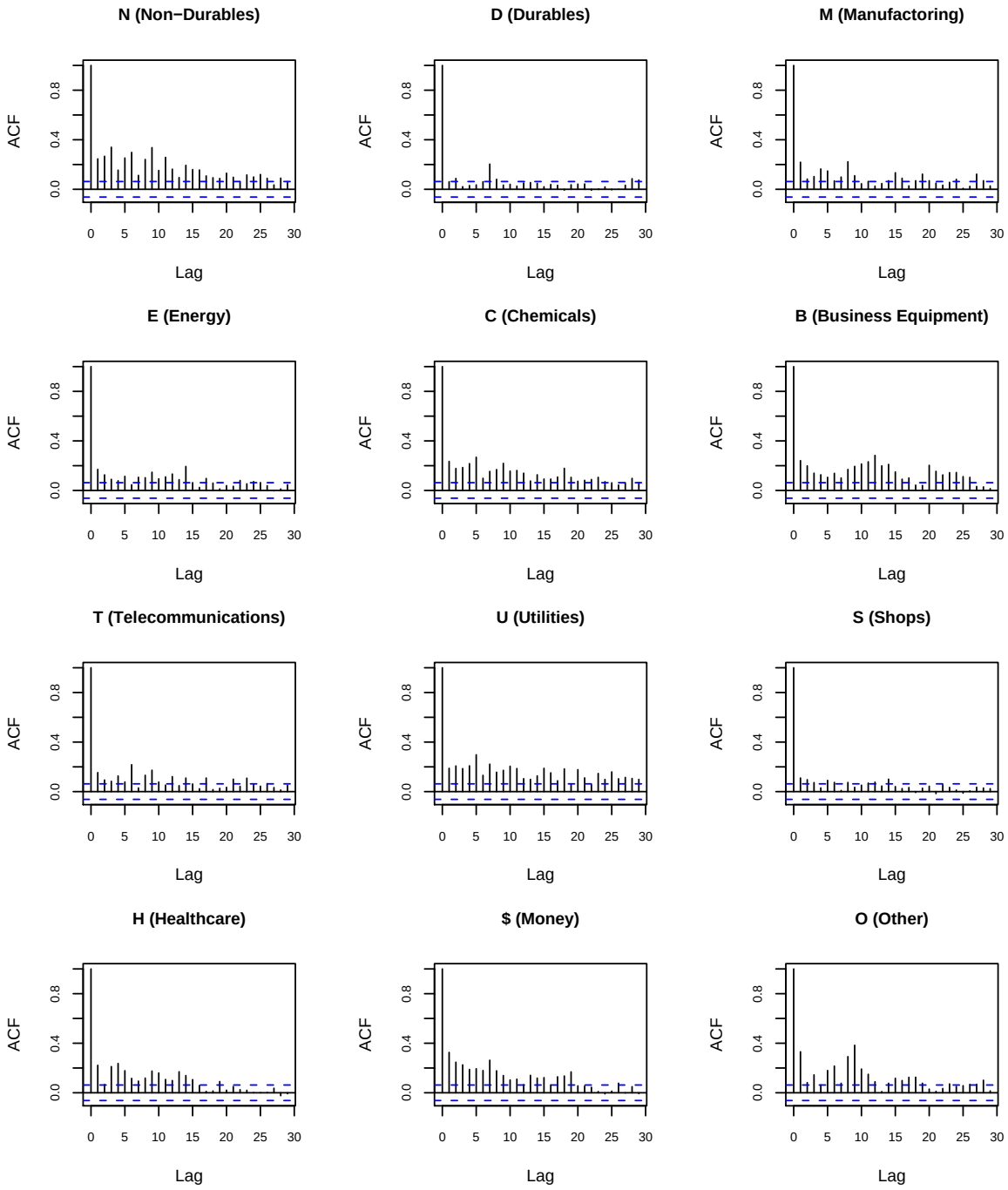

Figure 5.5: ACF plots of the squared standardized residuals $\left(\varepsilon_{ij}^{(mar)}\right)^2$ resulting from the marginal estimators $\boldsymbol{\beta}^{(mar)}$ and $\boldsymbol{\sigma}^{2(mar)}$

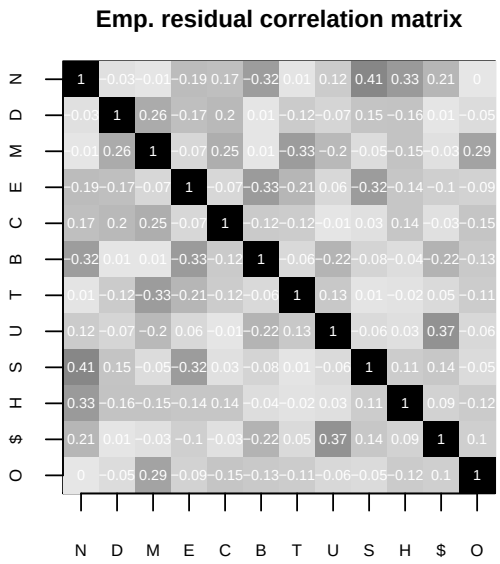

Figure 5.6: Visualization of the empirical residual correlation matrix $R^{(mar)}$

This suggests that one could also extend our model with fixed variance for all time points to a GARCH process.

Note that we have so far ignored the dependence between the portfolios besides the common covariate z_i. We will do this later in this chapter when we run MCMC algorithms. However, by using marginal estimates and the resulting residuals $\varepsilon_{ij}^{(mar)}$ for all $i \in \{1, \ldots, 978\}$, $j \in \{1, \ldots, 12\}$ we can calculate an empirical residual correlation matrix $R^{(mar)}$. We perform this with the R routine $cor()$.

We illustrate the estimated correlation matrix in the form of an image plot in Figure 5.6. Each entry in the image plots consists of a correlation and a color. The correlation belongs to the portfolios that can be found on the two axes, while a dark color marks a comparatively strong correlation (positive or negative) and a bright color marks a correlation close to 0. As we can see, there is no correlation whose absolute value is greater 0.5, which shows that a large fraction of the dependency between portfolio returns as observed in the time series plot can be explained by the market return. We see that from a total of 66 correlations, there are 7 with an absolute value greater 0.3 and 18 correlations between -0.05 and 0.05. The highest dependency can be observed between the residuals of portfolios Chemicals C and Consumer Non-Durables N, whose estimated correlation is 0.41.

5.2 Arranging the data in three different D-vines

To construct a D-vine structure for our partial correlation specification, we need to choose an order of the 12 Portfolio indices whose unordered set will be referred to as

$$\mathcal{I} = \{N, D, M, E, C, B, T, U, S, H, \$, O\}.$$

The chosen order of $\mathcal{I}$ then determines the first tree of the D-vine, which itself determines the other trees of the D-vine. As there are 12! possible orders of the portfolio index set and each reverse order leads to the same D-vine, there are $\frac{12!}{2} = 239500800$ different D-vines. We choose 3 of them.

5.2.1 1st Order: Original order

In the first order, we just keep the original order of the dataset. So in the first D-vine, the first tree is

$$N - D - M - E - C - B - T - U - S - H - \$ - O$$

5.2.2 2nd Order: Put strongest correlations on top

In Section 5.1, we used marginal estimators of $\boldsymbol{\beta}$ and $\boldsymbol{\sigma}^2$ to get a first estimator of the correlation matrix, which we called $R^{(mar)}$. We can use now these matrix entries to construct a D-vine, such that the edges of the first tree are correlations whose estimated absolute value is as large as possible. Then we expect that the absolute values of the estimated correlations from the MCMC algorithm in the first tree will be relatively large. In addition we hope that the partial correlations in the following trees may be close to 0 and therefore can be omitted in a reduced model.

We build the first tree of our D-vine by sequentially adding an edge that corresponds to the strongest possible of all correlations that are left without violating the D-vine rules. The lower triangle of our correlation matrix, as presented in Figure 5.6, is given in detail in Table 5.5. For readability, we omitted the upper triangle of this symmetric matrix.

The strongest correlations are $\rho_{SN} = 0.413$, $\rho_{\$U} = 0.367$, $\rho_{HN} = 0.332$, $\rho_{BE} = -0.328$, $\rho_{TM} = -0.325$ and $\rho_{SE} = -0.324$.

If we construct a graph of those correlations, such that every correlation corresponds to an edge of its indices, we get

$$\begin{gathered} B - E - S - N - H \\ \$ - U \\ T - M \end{gathered}$$

This graph consisting of three components is meant to be a part of the first tree of our D-vine.

Portf.	N	D	M	E	C	B	T	U	S	H	\$	O
N	1.00											
D	-0.03	1.00										
M	-0.01	0.26	1.00									
E	-0.19	-0.17	-0.07	1.00								
C	0.17	0.20	0.25	-0.07	1.00							
B	-0.32	0.01	0.01	**-0.33**	-0.12	1.00						
T	0.01	-0.12	**-0.33**	-0.21	-0.12	-0.06	1.00					
U	0.12	-0.07	-0.20	0.06	-0.01	-0.22	0.13	1.00				
S	**0.41**	0.15	-0.05	**-0.32**	0.03	-0.08	0.01	-0.06	1.00			
H	**0.33**	-0.16	-0.15	-0.14	0.14	-0.04	-0.02	0.03	0.11	1.00		
\$	0.21	0.01	-0.03	-0.10	-0.03	-0.22	0.05	**0.37**	0.14	0.09	1.00	
O	0.00	-0.05	0.29	-0.09	-0.15	-0.13	-0.11	-0.06	-0.05	-0.12	0.10	1.00

Table 5.5: Lower triangle of estimated correlation matrix $R^{(mar)}$; strongest empirical correlations are printed as bold values

Portf.	D	M	C	B	T	U	H	\$	O
D	1.00								
M	**0.26**	1.00							
C	0.20	**0.25**	1.00						
B	0.01	0.01	-0.12	1.00					
T	-0.12	(-0.33)	-0.12	-0.06	1.00				
U	-0.07	-0.20	-0.01	**-0.22**	0.13	1.00			
H	-0.16	-0.15	0.14	-0.04	-0.02	0.03	1.00		
\$	0.01	-0.03	-0.03	-0.22	0.05	(0.37)	0.09	1.00	
O	-0.05	**0.29**	-0.15	-0.13	-0.11	-0.06	-0.12	0.10	1.00

Table 5.6: Lower triangle of the submatrix of the estimated correlation matrix $R^{(mar)}$ without $\{E, S, N\}$; already used correlations are printed in brackets, the next strongest empirical correlations are printed as bold values

The next strongest correlation would be $\rho_{BN} = -0.320$, but if we add an edge $\{B, N\}$ to our graph, the vertex N would have three edges, which is not allowed for a D-vine. Since also the vertices E and S already have two edges, the search for the next strongest correlation is bounded on the submatrix of R shown in Table 5.6. All correlations that have already been used are printed in brackets (with E, S, N deleted).

The next strongest correlations in the submatrix are $\rho_{OM} = 0.294$, $\rho_{MD} = 0.260$, $\rho_{CM} = 0.246$ and $\rho_{\$B} = -0.223$. After adding an edge $\{MO\}$ to our graph, M is already connected to two vertices and therefore only another edge $\{\$B\}$ can be added to the graph. So far, we have expended our graph to

$$U - \$ - B - E - S - N - H$$
$$T - M - O$$

Now M, \$ and B are vertices with two edges, so we may remove all rows and columns of the submatrix of Table 5.6 corresponding to these three covariables, which means that we get a new submatrix of the correlation matrix as provided in Table 5.7. From that matrix, the strongest correlations are $\rho_{CD} = 0.197$ and $\rho_{HD} = -0.156$, so our graph expands to

$$U - \$ - B - E - S - N - H - D - C$$
$$T - M - O$$

Portf.	D	C	T	U	H	O
D	1.00					
C	**0.20**	1.00				
T	-0.12	-0.12	1.00			
U	-0.07	-0.01	0.13	1.00		
H	**-0.16**	0.14	-0.02	0.03	1.00	
O	-0.05	**-0.15**	-0.11	-0.06	-0.12	1.00

Table 5.7: Lower triangle of the submatrix of the estimated correlation matrix $R^{(mar)}$ without $\{E, S, N, M, \$, B\}$; the next strongest empirical correlations are printed as bold values

The next strongest correlation is $\rho_{OC} = -0.147$, so we can complete our graph into a tree by connecting the vertices C and O. This resulting tree is chosen as first tree of the second D-vine construction. It is

$$U - \$ - B - E - S - N - H - D - C - O - M - T$$

5.2.3 3rd Order: Sequentially choose minimal partial correlations

In the 3rd order, we try to find a D-vine which leads to zero partial correlations in the last trees of the vine. These partial correlations have a conditioned set with many covariables and therefore their corresponding likelihoods are harder to calculate by pair-copula algorithms. We again use the estimated correlation matrix $R^{(mar)}$ from the marginal estimates as basis for our choosing decisions, hoping that if some partial correlations derived by that empirical matrix are close to zero, this also holds for the corresponding partial correlations resulting from the MCMC algorithm. With this estimated covariance matrix, we are now able to calculate estimates for all partial correlations. We start by searching for covariables with indices $j, k \in \mathcal{I} = \{N, D, M, E, C, B, T, U, S, H, \$, O\}$, such that that the absolute value of the partial correlation

$$\hat{\rho}_{j,k|\mathcal{I}\setminus\{j,k\}}$$

is minimal.

We calculate all possible 66 partial correlations of that form and their absolute values. Table 5.8 shows the 26 lowest of them, and the remaining 40 partial correlations have an absolute value greater 0.1.

As the minimal absolute partial correlation is that of B and M given all others, we choose B as first node and M as last node of the first tree of the D-vine. This means the first tree is now

$$B - ... - M$$

and the last tree of the vine, i. e. the 11th tree, has been determined. The dots in the middle represent the remaining nodes and their adjacent edges of the first tree, which have not been assigned yet.

Rank	j	k	$\lvert\hat{\rho}_{j,k\lvert\mathcal{I}\setminus\{j,k\}}\rvert$
1	B	M	0.00027
2	E	U	0.00379
3	B	H	0.01430
4	D	\$	0.01882
5	T	\$	0.01925
6	N	O	0.01968
7	D	U	0.02035
8	C	U	0.02262
9	M	\$	0.02386
10	N	M	0.04257
11	T	U	0.04403
12	N	T	0.04895
13	N	\$	0.05131

Rank	j	k	$\lvert\hat{\rho}_{j,k\lvert\mathcal{I}\setminus\{j,k\}}\rvert$
14	H	\$	0.05283
15	U	H	0.05299
16	S	H	0.05875
17	N	U	0.06709
18	C	T	0.07256
19	S	\$	0.07301
20	U	O	0.07431
21	C	\$	0.07476
22	M	E	0.07858
23	\$	O	0.08121
24	C	S	0.08665
25	T	S	0.09485
26	N	C	0.09650

Table 5.8: All estimated partial correlations of the form $\hat{\rho}_{j,k|\mathcal{I}\setminus\{j,k\}}$ with absolute value less than 0.1, sorted by absolute value

Next, we want to assign covariables to the two next nodes of the first tree. We choose these covariables in a way that the sum of the absolute values of the partial correlations belonging to the 10th tree is minimal. That means we have to minimize the sum

$$|\hat{\rho}_{B,l|\mathcal{I}\setminus\{B,l,M\}}| + |\hat{\rho}_{m,M|\mathcal{I}\setminus\{m,M,B\}}|$$

The results of the lowest 10 possible sums are given in Table 5.9. So in our sense the best choice is H and \$, such that the first tree now looks like

$$B - \$ - ... - H - M$$

The next two covariates are chosen such that the two partial correlations determined by the two outer nodes of the 9th tree are as close to zero as possible, which means we have to minimize the sum

$$|\hat{\rho}_{B,l|\mathcal{I}\setminus\{B,l,M,H\}}| + |\hat{\rho}_{m,M|\mathcal{I}\setminus\{m,M,B,\$\}}| \tag{5.2}$$

Note that the inner partial correlation in the 9th tree is independent of the choice of l and m, but depends on on the indices that we have already assigned, so it does not have to be included in the minimization.

Table 5.10 shows the 5 combinations of partial correlations of the form (5.2) with lowest sum of absolute values.

Rank	l	$\hat{\rho}_{B,l\|\cdot}$	m	$\hat{\rho}_{m,M\|\cdot}$	$\lvert\hat{\rho}_{B,l\|\cdot}\rvert + \lvert\hat{\rho}_{m,M\|\cdot}\rvert$
1	H	-0.015	\$	0.024	0.039
2	H	-0.015	N	0.045	0.059
3	H	-0.015	E	-0.091	0.105
4	H	-0.015	S	-0.103	0.118
5	S	-0.102	\$	0.024	0.126
6	D	-0.110	\$	0.024	0.134
7	U	-0.121	\$	0.024	0.145
8	S	-0.102	N	0.045	0.147
9	D	-0.110	N	0.045	0.154
10	U	-0.121	N	0.045	0.166

Table 5.9: The 10 combinations of estimated partial correlations that are relevant for determining the 10th tree with lowest sum of absolute values

We see that the best combination is $\hat{\rho}_{B,S|\mathcal{I}\setminus\{B,S,M,H\}}$ and $\hat{\rho}_{N,M|\mathcal{I}\setminus\{N,M,B,\$\}}$, which is why we add N and S to the first tree of the vine, which looks now as follows:

$$B - \$ - N - ... - S - H - M$$

We go on with that scheme, always minimizing the sum of the absolute values of the two partial correlations determined by the two outer nodes of the current tree, as all partial correlations determined by inner points are independent of the next choices.

The next best results are shown in Table 5.11. These results suggest that we take as next covariables E and U, such that the first tree is now

$$B - \$ - N - E - ... - U - S - H - M$$

and the 8th tree has been determined.

A look at the results for the next step, as shown in Table 5.12, motivates us to add D and C to the first tree, which is now

$$B - \$ - N - E - D - ...- C - U - S - H - M$$

At last, only two possible choices are left, which are specified in Table 5.13. The best choice of the two leads finally to the complete first tree of the form

$$B - \$ - N - E - D - T - O - C - U - S - H - M$$

such that finally all trees of the D-vine are determined.

Rank	l	$\hat{\rho}_{B,l\|\cdot}$	m	$\hat{\rho}_{m,M\|\cdot}$	$\lvert\hat{\rho}_{B,l\|\cdot}\rvert + \lvert\hat{\rho}_{m,M\|\cdot}\rvert$
1	S	-0.10	N	0.047	0.15
2	D	-0.11	N	0.047	0.16
3	U	-0.12	N	0.047	0.17
4	S	-0.10	E	-0.092	0.19
5	D	-0.11	E	-0.092	0.20

Table 5.10: The 5 combinations of estimated partial correlations that are relevant for determining the 9th tree with lowest sum of absolute values

Rank	l	$\hat{\rho}_{B,l\|\cdot}$	m	$\hat{\rho}_{m,M\|\cdot}$	$\lvert\hat{\rho}_{B,l\|\cdot}\rvert + \lvert\hat{\rho}_{m,M\|\cdot}\rvert$
1	U	-0.11	E	-0.093	0.20
2	D	-0.12	E	-0.093	0.22
3	C	-0.16	E	-0.093	0.25
4	D	-0.12	U	-0.150	0.27
5	U	-0.11	D	0.178	0.29

Table 5.11: The 5 combinations of estimated partial correlations that are relevant for determining the 8th tree with lowest sum of absolute values

Rank	l	$\hat{\rho}_{B,l\|\cdot}$	m	$\hat{\rho}_{m,M\|\cdot}$	$\lvert\hat{\rho}_{B,l\|\cdot}\rvert + \lvert\hat{\rho}_{m,M\|\cdot}\rvert$
1	C	-0.16	D	0.20	0.35
2	D	-0.12	T	-0.26	0.37
3	D	-0.12	C	0.27	0.39
4	C	-0.16	T	-0.26	0.41
5	T	-0.23	D	0.20	0.43

Table 5.12: The 5 combinations of estimated partial correlations that are relevant for determining the 7th tree with lowest sum of absolute values

Rank	l	$\hat{\rho}_{B,l\|\cdot}$	m	$\hat{\rho}_{m,M\|\cdot}$	$\lvert\hat{\rho}_{B,l\|\cdot}\rvert + \lvert\hat{\rho}_{m,M\|\cdot}\rvert$
1	O	-0.21	T	-0.27	0.49
2	T	-0.21	O	0.30	0.52

Table 5.13: The two estimated partial correlations relevant for determining the 6th tree and their sum of absolute values

5.3 Results

We run our MCMC algorithm for each D-vine construction. As initial values for $\boldsymbol{\beta}$ and $\boldsymbol{\sigma}^2$ we take the marginal estimators that we have calculated in our data description. For the partial correlations, we use the empirical correlation matrix that we have calculated before and used to select the order in the 2nd and the 3rd D-vine construction. We use the definition of the partial correlation, but take our empirical correlation matrix instead of the true correlation matrix. Thus, we get estimates for every partial correlation in each D-vine construction and use them as initial values.

We choose the same prior parameters as in the simulations in the previous chapters, i.e. we take $s_j^2 := 100000 \quad \forall j \in \{1, \dots, 12\}$ for the prior variance of $\boldsymbol{\beta}$ and specify the prior parameters of σ_j^2 to $a_j := 1$ and $b_j = 0.001$ for every $j \in \{1, \dots, 12\}$. The prior distribution of the partial correlation parameters was already specificated in detail in the previous chapter and we leave it unchanged, that means we have a prior uniform distribution for the covariance matrix R which results from each partial correlation specification.

Again, we determine the proposal variance of each parameter by pilot runs, such that the acceptance rate is greater than 20% and resulting autocorrelations die down after up to 30 iterations. On the next pages, we discuss the results, which were calculated with 12000 MCMC iterations, a burn in period of 2000 and a subsample of the remaining chain consisting of every 30th iteration. This means that all estimates and quantities are based on 334 MCMC iterations.

At first, we compare the results for the marginal parameters $\boldsymbol{\beta}$ and $\boldsymbol{\sigma}^2$, as provided in Tables 5.14 to 5.16. In these tables, we use the following notation: θ denotes the current parameter of interest, whose estimates are provided in the corresponding row. $\hat{\theta}_{\alpha \cdot 100\%}$ denotes the empirical $\alpha \cdot 100\%$-quantile out of the 334 estimates for the parameter θ and for $\alpha \cdot 100\% \in \{2.5\%, 5\%, 95\%, 97.5\%\}$. $\hat{\theta}_{med}$ denotes the empirical median, $\bar{\theta}$ the mean and $\hat{\theta}_{mod}$ the estimated posterior mode of θ. The minimum of the subsample is denoted by $\hat{\theta}_{min}$, the maximum by $\hat{\theta}_{max}$. The estimator $\hat{\theta}_{IFM}$ corresponds to the marginal estimates $\boldsymbol{\beta}^{(mar)}$ and $\boldsymbol{\sigma}^{2(mar)}$ which we presented in Section 5.1 and used as initial values for $\boldsymbol{\beta}$ and $\boldsymbol{\sigma}^2$. At last, we added the mean acceptance rate $\bar{p}_{acc}$ which is 1 for all β_j parameters, since they are updated with a Gibbs sampler and not by an MH-step.

The results in Tables 5.14 to 5.16 are quite similar for all three vine orders. The mode estimates hardly differ, the greatest deviations can be observed on the estimates for high values of σ_j^2. For instance, the greatest absolute deviation between posterior mode estimates can be observed for the residual variance σ_H^2, where the estimate of the first construction is 11.719, of the second 12.001 and of the third 11.905. The corresponding marginal estimator is 11.803. However, there is no case where the estimator $\hat{\theta}_{mod}$ of one construction is outside the interval $(\hat{\theta}_{5\%}, \hat{\theta}_{95\%})$ of another constrution. We also observe that the differences between the posterior modes and the marginal estimators denoted in the IFM-column are quite small. To conclude, the mean acceptance rate of the σ^2-Parameters is about 30%.

θ	$\hat{\theta}_{min}$	$\hat{\theta}_{2.5\%}$	$\hat{\theta}_{5\%}$	$\hat{\theta}_{med}$	$\bar{\theta}$	$\hat{\theta}_{95\%}$	$\hat{\theta}_{97.5\%}$	$\hat{\theta}_{max}$	$\hat{\theta}_{mod}$	$\hat{\theta}_{IFM}$	$\bar{p}_{acc}$
β_N	0.733	0.746	0.748	0.768	0.768	0.788	0.793	0.802	0.768	0.767	1.000
β_D	1.157	1.169	1.178	1.219	1.217	1.253	1.258	1.273	1.220	1.218	1.000
β_M	1.152	1.168	1.171	1.188	1.189	1.209	1.212	1.224	1.187	1.189	1.000
β_E	0.807	0.817	0.822	0.859	0.859	0.899	0.905	0.943	0.860	0.858	1.000
β_C	0.935	0.946	0.952	0.979	0.979	1.005	1.008	1.025	0.979	0.978	1.000
β_B	1.237	1.254	1.261	1.291	1.292	1.325	1.329	1.336	1.291	1.294	1.000
β_T	0.584	0.607	0.614	0.640	0.642	0.673	0.679	0.692	0.638	0.643	1.000
β_U	0.756	0.767	0.771	0.806	0.806	0.841	0.850	0.862	0.807	0.805	1.000
β_S	0.916	0.926	0.932	0.960	0.960	0.987	0.994	1.013	0.960	0.961	1.000
β_H	0.820	0.833	0.835	0.866	0.866	0.900	0.903	0.927	0.864	0.865	1.000
$\beta_\$$	1.107	1.124	1.132	1.158	1.158	1.183	1.188	1.208	1.157	1.158	1.000
β_O	1.086	1.108	1.112	1.135	1.136	1.161	1.168	1.177	1.135	1.135	1.000
σ_N^2	4.442	4.618	4.680	5.046	5.047	5.417	5.476	5.671	5.051	5.017	0.286
σ_D^2	12.747	13.339	13.516	14.550	14.571	15.587	15.882	16.313	14.479	14.427	0.311
σ_M^2	3.525	3.775	3.807	4.107	4.118	4.449	4.495	4.652	4.097	4.083	0.283
σ_E^2	12.668	13.263	13.376	14.392	14.437	15.541	15.675	16.158	14.407	14.412	0.279
σ_C^2	5.908	6.033	6.126	6.586	6.592	7.090	7.133	7.566	6.578	6.548	0.304
σ_B^2	9.203	9.725	9.855	10.581	10.577	11.292	11.509	12.023	10.601	10.521	0.285
σ_T^2	8.075	8.279	8.353	8.997	9.025	9.757	9.813	10.363	8.961	8.942	0.300
σ_U^2	12.082	12.424	12.563	13.527	13.558	14.667	14.874	15.441	13.512	13.456	0.295
σ_S^2	6.824	7.142	7.268	7.907	7.891	8.496	8.584	9.076	7.926	7.829	0.292
σ_H^2	10.596	10.835	10.981	11.784	11.837	12.892	13.038	13.514	11.719	11.803	0.302
$\sigma_\$^2$	6.685	6.943	7.038	7.608	7.584	8.146	8.271	8.592	7.617	7.527	0.305
σ_O^2	6.134	6.292	6.336	6.804	6.830	7.393	7.492	7.943	6.773	6.790	0.300

Table 5.14: MCMC results for the marginal parameters of the 1st vine construction.

θ	$\hat{\theta}_{min}$	$\hat{\theta}_{2.5\%}$	$\hat{\theta}_{5\%}$	$\hat{\theta}_{med}$	$\bar{\theta}$	$\hat{\theta}_{95\%}$	$\hat{\theta}_{97.5\%}$	$\hat{\theta}_{max}$	$\hat{\theta}_{mod}$	$\hat{\theta}_{IFM}$	$\bar{p}_{acc}$
β_N	0.728	0.741	0.745	0.769	0.768	0.787	0.789	0.798	0.771	0.767	1.000
β_D	1.166	1.179	1.185	1.219	1.220	1.256	1.262	1.302	1.219	1.218	1.000
β_M	1.142	1.165	1.168	1.190	1.190	1.209	1.213	1.222	1.190	1.189	1.000
β_E	0.800	0.815	0.823	0.860	0.859	0.894	0.899	0.915	0.862	0.858	1.000
β_C	0.941	0.949	0.952	0.980	0.979	1.002	1.009	1.032	0.980	0.978	1.000
β_B	1.235	1.254	1.258	1.291	1.292	1.324	1.333	1.366	1.289	1.294	1.000
β_T	0.593	0.611	0.614	0.645	0.644	0.671	0.675	0.701	0.645	0.643	1.000
β_U	0.754	0.765	0.769	0.804	0.803	0.835	0.842	0.864	0.804	0.805	1.000
β_S	0.905	0.931	0.936	0.964	0.963	0.988	0.991	1.010	0.964	0.961	1.000
β_H	0.808	0.829	0.835	0.866	0.866	0.899	0.903	0.915	0.866	0.865	1.000
$\beta_\$$	1.109	1.128	1.132	1.159	1.159	1.184	1.189	1.219	1.159	1.158	1.000
β_O	1.092	1.105	1.108	1.133	1.134	1.161	1.166	1.178	1.134	1.135	1.000
σ_N^2	4.530	4.619	4.710	5.044	5.047	5.406	5.545	5.829	5.042	5.017	0.281
σ_D^2	12.642	13.246	13.453	14.401	14.481	15.571	15.945	16.872	14.356	14.427	0.296
σ_M^2	3.596	3.764	3.829	4.120	4.117	4.391	4.451	4.664	4.119	4.083	0.290
σ_E^2	12.595	13.237	13.488	14.372	14.427	15.518	15.715	16.366	14.311	14.412	0.280
σ_C^2	5.845	6.065	6.138	6.606	6.596	7.066	7.142	7.311	6.608	6.548	0.294
σ_B^2	9.418	9.807	9.879	10.616	10.637	11.498	11.642	12.779	10.587	10.521	0.284
σ_T^2	7.938	8.206	8.311	9.019	9.017	9.682	9.838	10.392	9.034	8.942	0.297
σ_U^2	11.758	12.257	12.512	13.522	13.548	14.525	14.991	15.750	13.483	13.456	0.298
σ_S^2	7.027	7.251	7.337	7.874	7.881	8.529	8.597	9.003	7.858	7.829	0.287
σ_H^2	10.602	10.957	11.024	11.968	11.931	12.830	13.020	13.465	12.001	11.803	0.301
$\sigma_\$^2$	6.593	6.951	7.044	7.538	7.564	8.150	8.263	8.737	7.524	7.527	0.301
σ_O^2	6.128	6.286	6.341	6.846	6.845	7.342	7.487	7.800	6.857	6.790	0.293

Table 5.15: MCMC results for the marginal parameters of the 2nd vine construction.

θ	$\hat{\theta}_{min}$	$\hat{\theta}_{2.5\%}$	$\hat{\theta}_{5\%}$	$\hat{\theta}_{med}$	$\bar{\theta}$	$\hat{\theta}_{95\%}$	$\hat{\theta}_{97.5\%}$	$\hat{\theta}_{max}$	$\hat{\theta}_{mod}$	$\hat{\theta}_{IFM}$	$\bar{p}_{acc}$
β_N	0.729	0.741	0.746	0.767	0.768	0.789	0.792	0.813	0.767	0.767	1.000
β_D	1.136	1.173	1.183	1.221	1.219	1.254	1.261	1.296	1.221	1.218	1.000
β_M	1.159	1.169	1.171	1.190	1.189	1.208	1.210	1.223	1.189	1.189	1.000
β_E	0.793	0.815	0.823	0.859	0.859	0.894	0.898	0.917	0.859	0.858	1.000
β_C	0.934	0.951	0.955	0.979	0.979	1.005	1.013	1.019	0.979	0.978	1.000
β_B	1.218	1.258	1.263	1.295	1.294	1.326	1.333	1.352	1.295	1.294	1.000
β_T	0.594	0.606	0.614	0.644	0.643	0.670	0.678	0.691	0.645	0.643	1.000
β_U	0.744	0.758	0.765	0.807	0.806	0.839	0.848	0.867	0.809	0.805	1.000
β_S	0.910	0.929	0.935	0.961	0.962	0.992	0.996	1.018	0.960	0.961	1.000
β_H	0.806	0.825	0.828	0.865	0.866	0.904	0.910	0.928	0.866	0.865	1.000
$\beta_\$$	1.117	1.127	1.130	1.157	1.157	1.185	1.189	1.200	1.156	1.158	1.000
β_O	1.089	1.105	1.110	1.133	1.134	1.160	1.164	1.170	1.132	1.135	1.000
σ_N^2	4.344	4.551	4.649	5.012	5.024	5.449	5.526	5.788	4.979	5.017	0.277
σ_D^2	12.676	13.276	13.421	14.549	14.538	15.649	16.116	16.823	14.554	14.427	0.309
σ_M^2	3.456	3.766	3.817	4.089	4.096	4.395	4.451	4.801	4.069	4.083	0.281
σ_E^2	12.651	13.330	13.538	14.431	14.500	15.681	15.876	16.368	14.364	14.412	0.273
σ_C^2	5.837	6.115	6.171	6.616	6.622	7.133	7.217	7.591	6.625	6.548	0.300
σ_B^2	9.384	9.715	9.815	10.611	10.617	11.478	11.607	11.920	10.581	10.521	0.280
σ_T^2	7.954	8.171	8.327	9.008	9.004	9.652	9.850	10.147	9.020	8.942	0.296
σ_U^2	12.039	12.346	12.521	13.536	13.546	14.591	14.689	15.312	13.565	13.456	0.292
σ_S^2	6.778	7.200	7.316	7.842	7.863	8.453	8.543	9.038	7.805	7.829	0.289
σ_H^2	10.606	10.810	10.935	11.874	11.890	12.747	13.006	13.866	11.905	11.803	0.293
$\sigma_\$^2$	6.633	6.946	7.062	7.558	7.584	8.206	8.388	8.756	7.532	7.527	0.293
σ_O^2	6.195	6.331	6.366	6.820	6.827	7.298	7.367	7.723	6.839	6.790	0.293

Table 5.16: MCMC results for the marginal parameters of the 3rd vine construction.

To visualize the results for the partial correlation parameters, we use image plots that are constructed as follows: Each square of the plot shows the value of a certain (partial) correlation between the returns of the two portfolios that belong to its coordinates, let's say A on the horizontal axis and B on the vertical axis. The corresponding conditioning set is formed by all portfolio indices that lie on the vertical axis between A and B.
Of course, if A is adjacent to B on that axis, the conditioning set is empty, and therefore we have the product moment correlation ρ_{AB}. If we look at another square e. g. with coordinates A and D, and if we see that portfolio indices B and C lie between A and D on the vertical axis, than we know that this square represents the partial correlation $\rho_{AD|BC}$.

This means that our image plots for the partial correlations are constructed in such a way that the diagonal is empty and each subdiagonal below shows all partial correlations of one specific tree of the vine: The first subdiagonal shows the unconditional correlations of the first tree of the vine, the second subdiagonal all partial correlations on the second tree, and so on.

To see how these plots are constructed we start with an example: Consider a four dimensional vine whose first tree is

$$A - B - C - D$$

We set $\rho_{AB} := \rho_{BC} := \rho_{CD} = 0.8$, $\rho_{AC|B} := \rho_{BD|C} := 0.5$ and $\rho_{AD|BC} = 0.3$. Then the corresponding image plot is provided in Figure 5.7, and the location of each parameter in the plot can be looked up in the table right next to the plot.

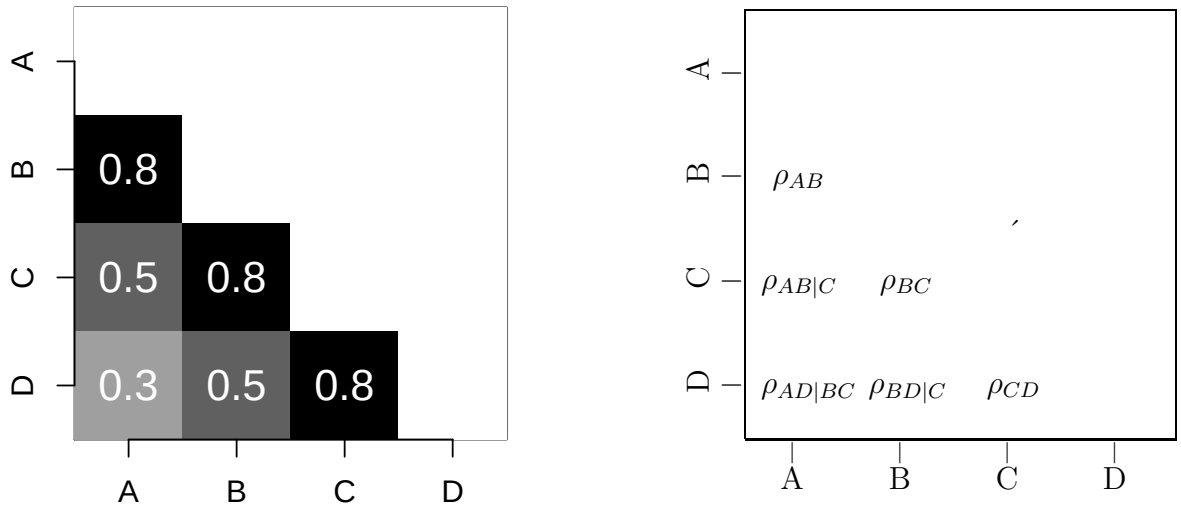

Figure 5.7: Image plot of the discussed example and position of the partial correlations

In Figure 5.8, we see the posterior mode estimates for each partial correlation and each order in form of image plots that are constructed as explained before. More detailed information on the results is provided in the appendix.

As we have different partial correlation specifications for each vine, it is difficult to compare the estimated values. But as every construction results from a multivariate normal distribution, we expect that the distribution resulting from the estimates should be – at least approximately – the same for the three orders. However, we can derive an estimated correlation matrix from the mode estimators for the partial correlations for each vine. We then expect that the estimates of the correlation matrix are quite similar.
As we can see in Figure 5.9, this is the case.

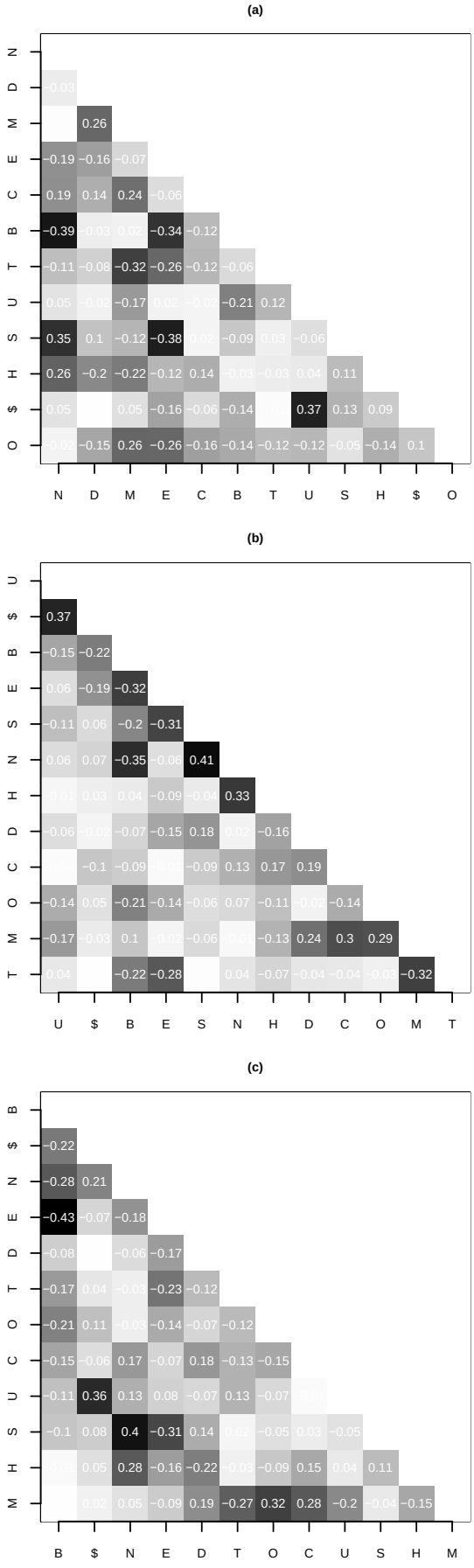

Figure 5.8: Estimated partial correlations resulting from the mode estimates in the 1st (a), the 2nd (b) and the 3rd (c) order

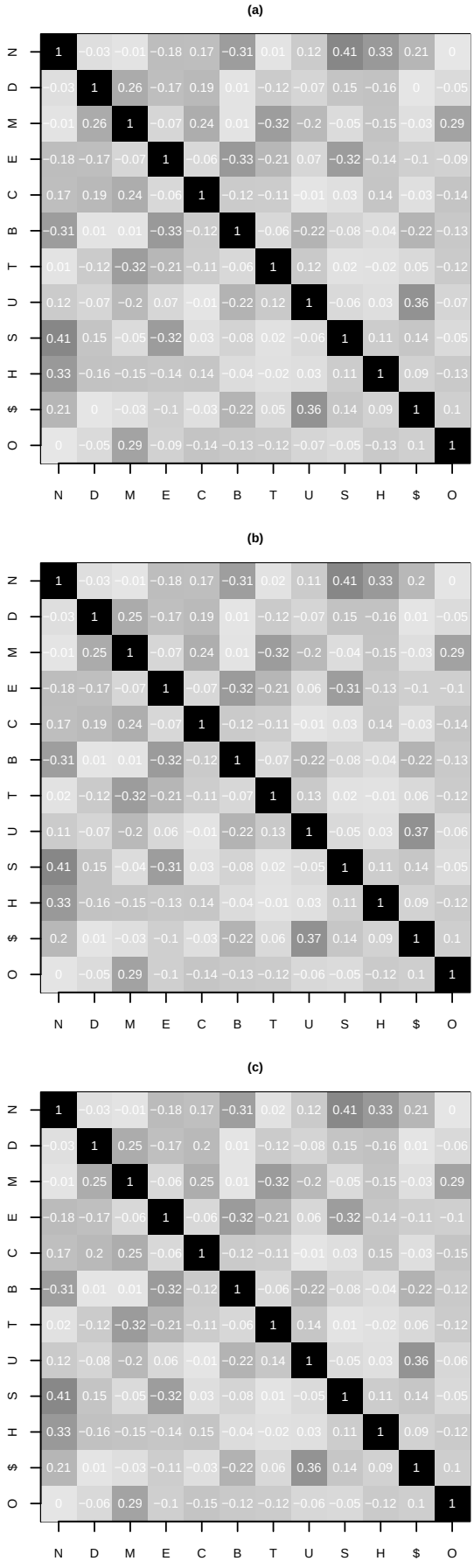

Figure 5.9: Estimated Correlation matrices resulting from the mode estimates for the partial correlations in the 1st (a), the 2nd (b) and the 3rd (c) order

5.4 Model reduction and comparison with full models

As there are 24 marginal and 66 copula parameters in our three models, the question arises whether it is possible to omit some of them, e.g. fix those parameters to zero. To decide this, we construct estimated credible intervals for each parameter $\theta \in \{\beta_j, \sigma_j^2, \rho_{jk|j+1:k-1},\ j = 1, \ldots, 12,\ k = j+1, \ldots, 12\}$ from our MCMC sample. If $0 \in (\hat{\theta}_{5\%}, \hat{\theta}_{95\%})$ holds, we know that θ is not credible on the level 10% based on our estimation. Similarly, we infer that θ is not credible on the level 5% if $0 \in (\hat{\theta}_{2.5\%}, \hat{\theta}_{97.5\%})$. Since the second credible interval always contains the first, it is clear that θ is not credible on the 5% level if it is not credible on the 10% level.

Looking in Tables 5.14 to 5.16 we see that no interval $(\hat{\theta}_{2.5\%}, \hat{\theta}_{97.5\%})$ contains 0, thus all β_j parameters are credible on the 5% level in every D-vine construction. So we have no motivation to leave out some β_j components in a reduced model, which would induce independency of the associated portfolio return from the market return. For the σ_j^2 parameters, a value of 0 makes no sense and is also not allowed according to our premises. Instead, one could think of choosing the same σ^2 parameter for all the portfolio residual variances σ_j^2, which would reduce the number of parameters by 11. However, the Tables 5.14 to 5.16 suggest that most variances are different, since for instance the sampled maximum of σ_N^2 is smaller than the sampled minimum of σ_D^2 in all three orders.

So we concentrate in the following on reducing models by leaving out all partial correlation parameters that are not credible on the 10% level resp. on the 5% level, such that we get two reduced models for each D-vine construction. For the first order, we see that 17 partial correlation parameters are not credible on the 10% level, and another three not credible on the 5% level. So almost one third of the copula parameters may be left out in a reduced model.

For the second construction, we expect that all unconditional correlations, i. e. those in the top tree, are credible on both levels. Looking at the results, we see that this is the case. On the other hand, we hoped that many of the partial correlations in the other trees would be zero. Indeed, 19 resp. 23 partial correlation parameters may be omitted in the reduced model if one integrates only the parameters that are credible on the 10% resp. on the 5% level. This is more than in any other examined order.

We pursued another plan in the 3rd construction: Here, the partial correlations with the largest conditioning set should be non-credible, if possible. From the results we see that this is the case for 5 of the 6 partial correlations in the three bottom trees on both considered credible levels. In total, there are 16 non-credible partial correlations based on the 10% level and 20 based on the 5% level.

We run our MCMC algorithm on all mentioned reduced models by leaving out the updates of all parameters that were fixed to 0. The results for the partial correlation parameters are shown in Figures 5.10 and 5.11. On the left hand side of the figure are the image plots representing the estimates of the partial correlations. To compare each reduced model to the corresponding full model, one can look at the right hand side, where the difference of each parameter between reduced and full model is shown. As one can see, the estimates of the reduced models are very similar to those of the corresponding full models. This is confirmed by the look on the difference shown in the right columns of the two figures.

But of course, one is again interested at the correlation matrices resulting from the partial correlation estimates. Because of the small difference between the partial correlations in the full and the reduced models, one expects that also the correlation matrices look very similar. To check this, one can look at Figures 5.12 and 5.13, where the estimated correlation matrices from the reduced models and the difference to those of the full models are shown as image plot for each D-vine construction and each credibility level.

To compare both full and reduced models, we use following Spiegelhalter et al. (2002) and Silva and Lopes (2008) the model decision criterion DIC, which is defined as

$$DIC = 2E\left(D\left(\boldsymbol{\theta}\right)|\boldsymbol{y}\right) - D\left(E(\boldsymbol{\theta}|\boldsymbol{y})\right)$$

where $D(\cdot)$ denotes the deviance which is defined as $D(\boldsymbol{\theta}) := -2\log\left(f(\boldsymbol{y}|\boldsymbol{\theta})\right)$. A model with a smaller DIC is preferred to a model with a larger DIC. With a (sub-)sample $\{\boldsymbol{\theta}^{(r)},\ r = 1, \ldots, s\}$ of size s generated by the MCMC algorithm, one can estimate the DIC by

$$\widehat{DIC} := 2\frac{1}{s}\sum_{r=1}^{s} D(\boldsymbol{\theta}^{(r)}) - D\left(\frac{1}{s}\sum_{r=1}^{s}\boldsymbol{\theta}^{(r)}\right)$$

The results are provided in Table 5.17. Over all models, the DIC method prefers the 2nd order model reduced by all parameters that are not credible on the level 10%. Regarding all full models, the 1st order is preferred. The model with the smallest number of parameters is the reduced model of the 2nd D-vine construction based on the level 5%.

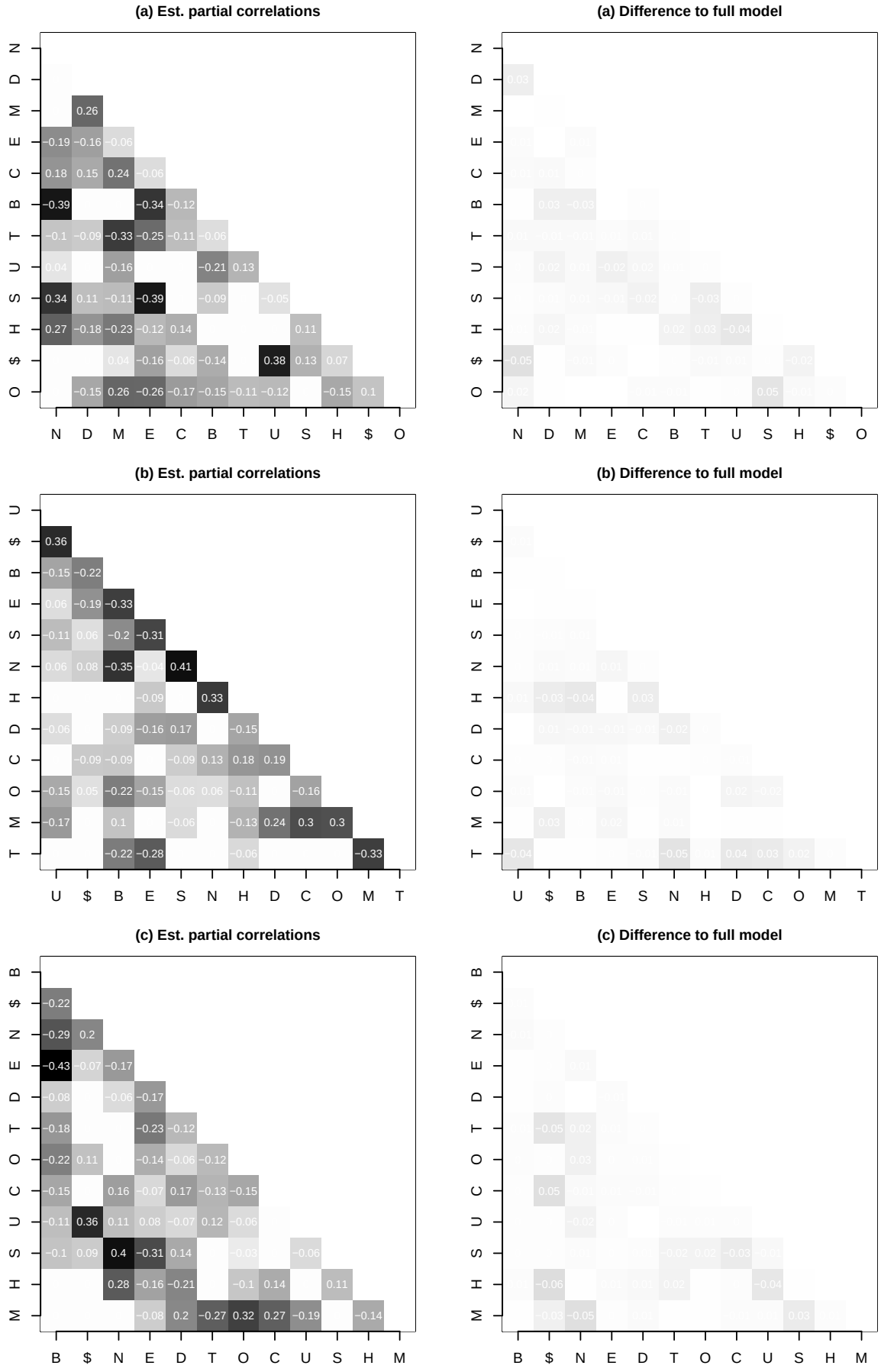

Figure 5.10: Estimated partial correlations of the reduced models and their difference to the corresponding full models in the 1st (a), the 2nd (b) and the 3rd (c) order. Credible Level: 10%

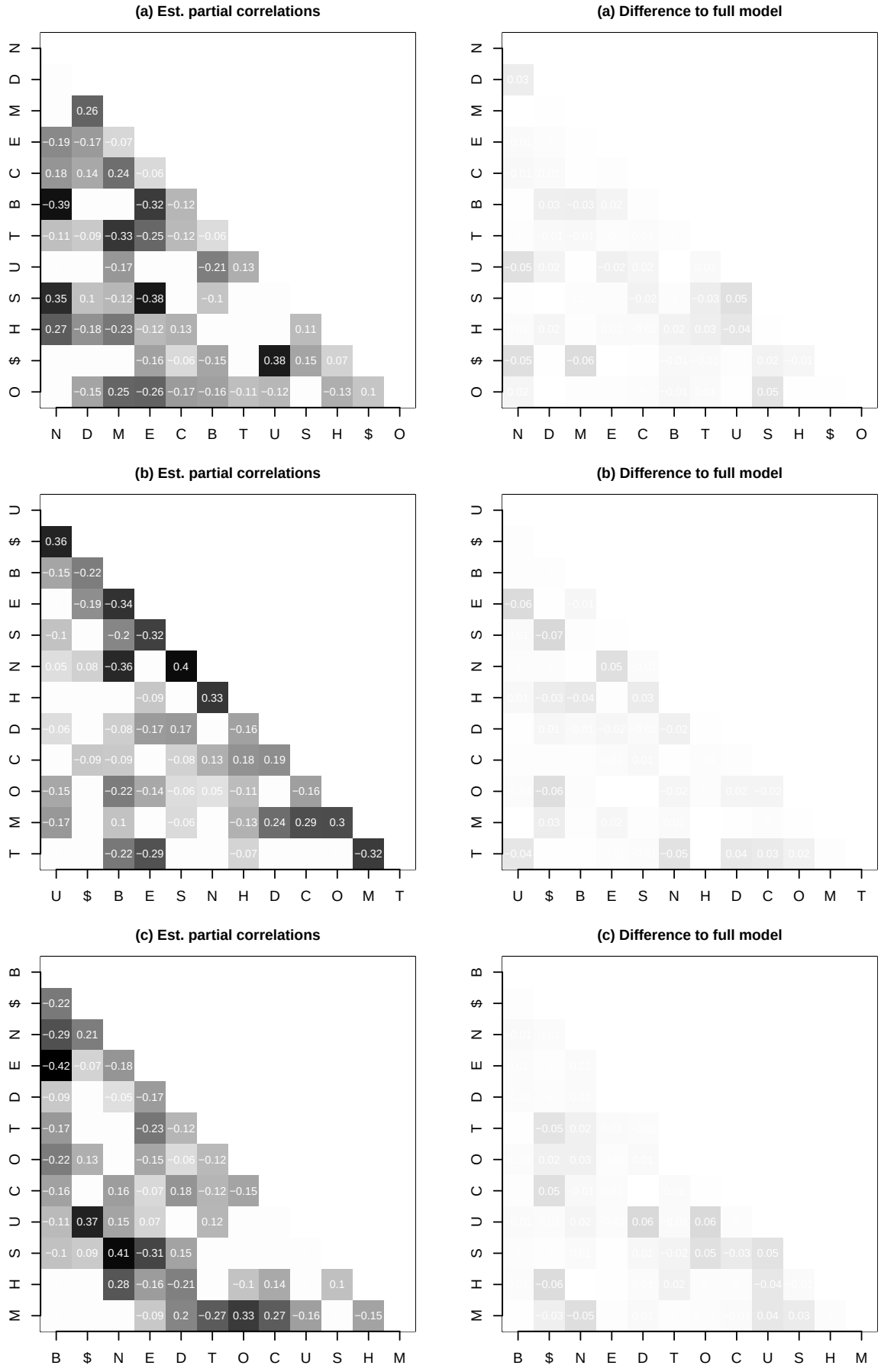

Figure 5.11: Estimated partial correlations of the reduced models and their difference to the corresponding full models in the 1st (a), the 2nd (b) and the 3rd (c) order. Credible Level: 5%

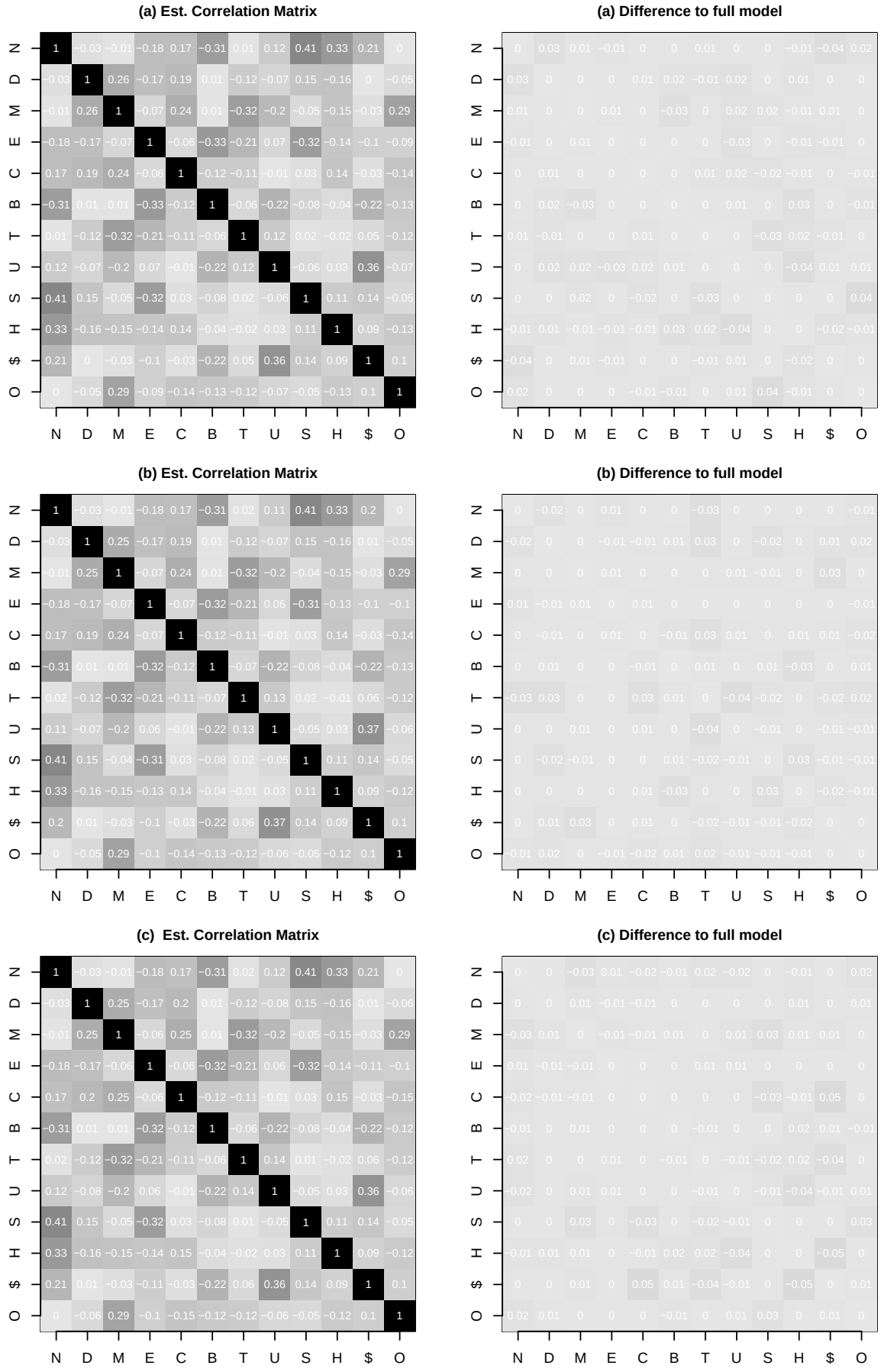

Figure 5.12: Estimated correlation matrices of the reduced models and their difference to the corresponding full models in the 1st (a), the 2nd (b) and the 3rd (c) order. Credible Level: 10%

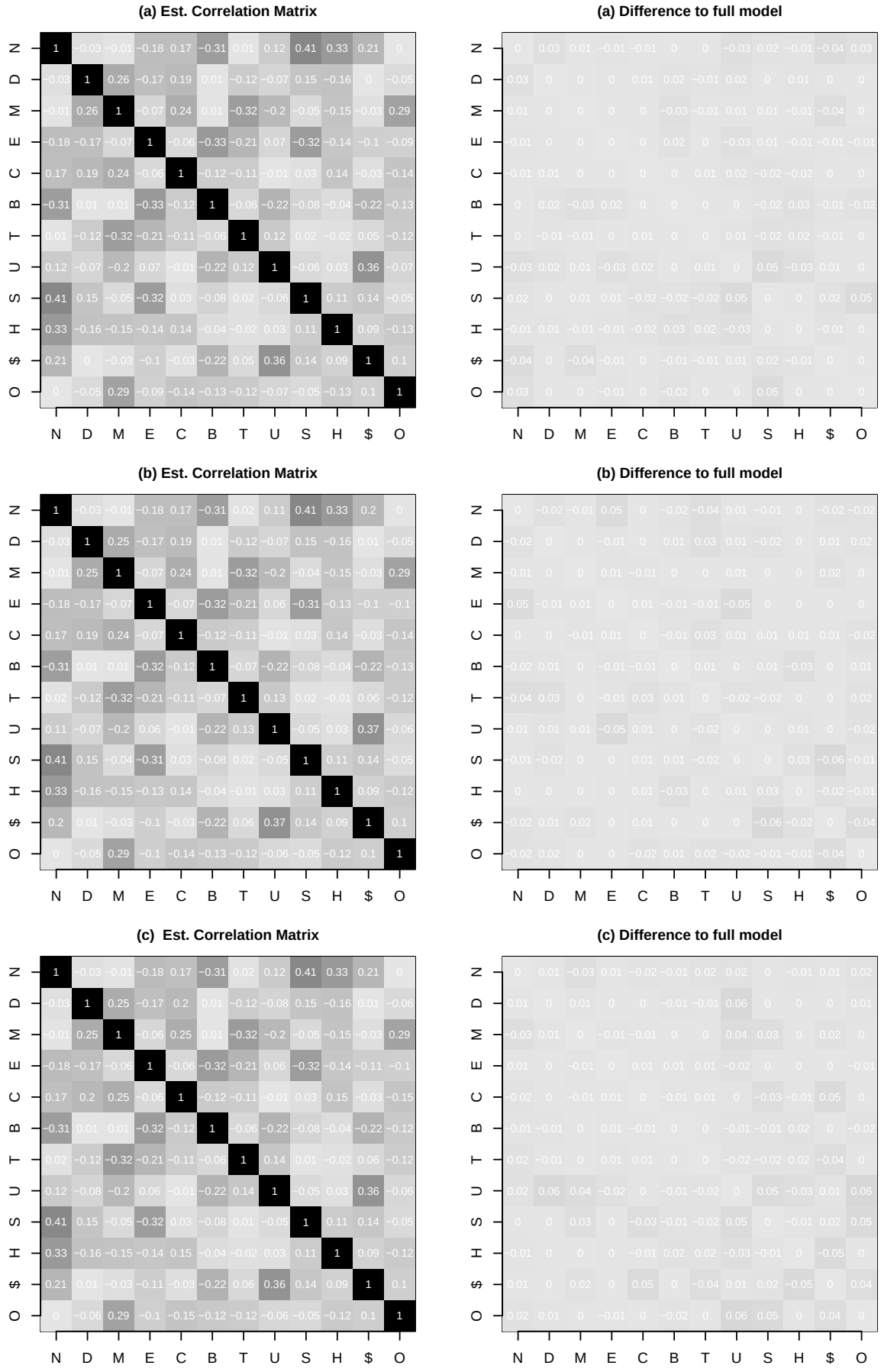

Figure 5.13: Estimated correlation matrices of the reduced models and their difference to the corresponding full models in the 1st (a), the 2nd (b) and the 3rd (c) order. Credible Level: 5%

Model	est. DIC	# Param.
1st Order	56849.98	80
Reduced (10%) 1st Order	56831.41	63
Reduced (5%) 1st Order	56837.28	60
2nd Order	56851.12	80
Reduced (10%) 2nd Order	56825.17	61
Reduced (5%) 2nd Order	56832.25	57
3rd Order	56850.63	80
Reduced (10%) 3rd Order	56834.56	64
Reduced (5%) 3rd Order	56834.56	60

Table 5.17: Comparison of reduced and full models for all three D-vine constructions

5.5 Model validation

We are now interested in the ability of our model to predict portfolio returns if the market excess return is given. For that purpose, we construct statistics consisting of proportions of each of the 12 industrial returns. Representatively, we choose 5 statistics $S_{i1}, \dots, S_{i5}$ for each point in time $i \in \{1, \dots, 978\}$, that are given by

$$S_{ik} = \sum_{j=1}^{12} w_{kj} Y_{ij}$$

where the weights w_{kj} sum up to 1 and are constant over time for all $k \in \{1, \dots, 5\}$. The weights of the five statistics are defined in Table 5.18 and are based on the empirical estimates of the correlation matrix $R = (\rho_{jk})_{j,k=1,\dots,12}$ and the residual variances $\sigma_1^2, \dots, \sigma_{12}^2$. For simplicity, we use the notations $\hat{\rho}_{jk} := \rho_{jk}^{(mar)}$ and $\hat{\sigma_j} := \sqrt{\sigma_j^{2(mar)}} \quad (j, k = 1, \dots, 12)$ in the table. A visualization of the resulting weights for each statistic is provided in Figure 5.14.

In the first statistic S_{i1}, we choose equal weights for each portfolio. The weights of the second statistic for each portfolio $j \in \{1, \dots, 12\}$ are the higher the stronger the portfolio j is correlated to all other portfolios. For the third statistic, we modify this choice by using only positive correlations and ignoring all negative correlations. This is due to the fact that a positive correlation between two assets leads to a higher risk, whereas a negative correlation reduces the risk. For the weights of statistic S_{i4}, we restrict only on correlations that are at least 0.1. The last statistic S_{i5} also takes into account the residual variances. Here, we choose the weights based on all positive covariances, including the residual variance itself. For each portfolio $j \in \{1, \dots, 12\}$, the weight w_{5j} is proportional to the sum of the variance σ_j^2 and the positive covariances between Y_{ij} and all other portfolio returns.

We now omit the years 2001 to 2007 in our data, that is all excess returns with time indices starting from 895, and want to apply the CAPM model on the reduced data set.

Statistic	Definition of weights	
S_{i1}	$w_{1j} = \frac{1}{12}$	$(j = 1, \ldots, 12)$
S_{i2}	$w_{2j} := \frac{a_j}{\sum_{k=1}^{12} a_{2k}}$ with $a_{2j} := \sum_{\substack{k=1 \\ k \neq j}}^{12} \lvert\hat{\rho}_{jk}\rvert$	$(j = 1, \ldots, 12)$
S_{i3}	$w_{3j} := \frac{a_j}{\sum_{k=1}^{12} a_{3k}}$ with $a_{3j} := \sum_{\substack{k=1 \\ k \neq j}}^{12} \hat{\rho}_{jk} \mathbf{1}_{\{\hat{\rho}_{jk} > 0\}}$	$(j = 1, \ldots, 12)$
S_{i4}	$w_{4j} := \frac{a_j}{\sum_{k=1}^{12} a_{4k}}$ with $a_{4j} := \sum_{\substack{k=1 \\ k \neq j}}^{12} \hat{\rho}_{jk} \mathbf{1}_{\{\hat{\rho}_{jk} > 0.1\}}$	$(j = 1, \ldots, 12)$
S_{i5}	$w_{5j} := \frac{a_j}{\sum_{k=1}^{12} a_{4k}}$ with $a_{5j} := \sum_{k=1}^{12} \hat{\sigma}_j \hat{\sigma}_k \hat{\rho}_{jk} \mathbf{1}_{\{\hat{\rho}_{jk} > 0\}}$	$(j = 1, \ldots, 12)$

Table 5.18: Definition of weights for statistics $S_{i1}, \ldots, S_{i5}$

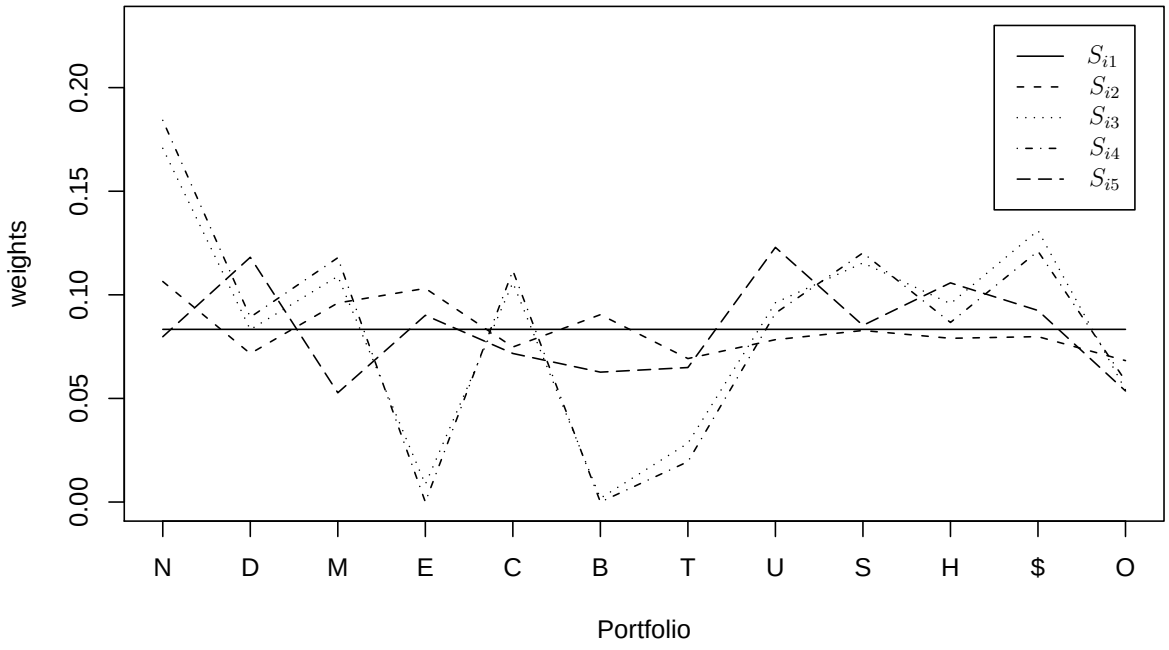

Figure 5.14: Illustration of weights for the statistics $S_{i1}, \ldots, S_{i5}$

In Section 5.4 we have seen that the DIC method prefers the model based on the 2nd D-vine construction reduced by all parameters that are not credible on the 10% level. So we choose this model and run our MCMC algorithm.

We run this algorithm until iteration $m = 52000$, choose a burn-in period of 2000 and use every 10th iteration. This means we get a sample

$$\{\boldsymbol{\theta}^{(r)}, \quad r = 1, \dots, 5000\}$$

of the posterior distribution of the parameter vector $\boldsymbol{\theta} = \left(\boldsymbol{\beta}', \boldsymbol{\sigma}^{2\prime}, \boldsymbol{\rho}_{\mathcal{V}}'\right)'$ where $\boldsymbol{\rho}_{\mathcal{V}}$ denotes the parameter vector of partial correlations in the considered model.

For each remaining dates $i \in \{895, \dots, 978\}$ that have not been used in our previous calculation, we are interested in the predictive quantiles of our statistics $S_{ik},\ k \in \{1, \dots, 5\}$ given the market return:

$$q_{ik}^{\alpha\cdot 100\%} := \inf_{x \in \mathbb{R}} \left\{ P(S_{ik} \leq x | \boldsymbol{y}_1, \dots, \boldsymbol{y}_{i-1}, z_1, \dots, z_i) \geq \alpha \right\}$$

To estimate these quantiles, we look at the situation where also $\boldsymbol{\theta}$ is given. At first, we calculate the quantiles

$$\begin{aligned} q_{ik}^{\alpha\cdot 100\%}(\boldsymbol{\theta}) &:= \inf_{x \in \mathbb{R}} \left\{ P(S_{ik} \leq x | \boldsymbol{y}_1, \dots, \boldsymbol{y}_{i-1}, z_1, \dots, z_i, \boldsymbol{\theta}) \geq \alpha \right\} \\ &= \inf_{x \in \mathbb{R}} \left\{ P(S_{ik} \leq x | z_i, \boldsymbol{\theta}) \geq \alpha \right\} \end{aligned}$$

From the model definition (4.1), we know that the distribution of $\boldsymbol{Y}_i = (Y_{i,1}, \dots, Y_{i,12})'$ for given $\boldsymbol{\theta}$ is

$$\boldsymbol{Y}_i | \boldsymbol{\theta} \sim \mathcal{N}\left(z_i \boldsymbol{\beta}, \Sigma\right)$$

where $\Sigma = \operatorname{diag}(\boldsymbol{\sigma}) R \operatorname{diag}(\boldsymbol{\sigma})$ and R denotes the correlation matrix resulting from $\boldsymbol{\rho}_{\mathcal{V}}$. Since $S_{ik} = \sum_{i=1}^{12} w_{kj} Y_{ij} = \boldsymbol{w}_k' \boldsymbol{Y}_i$ for $\boldsymbol{w}_k := (w_{1,k}, \dots, w_{12,k})'$, it holds

$$S_{ik} | \boldsymbol{\theta} \sim \mathcal{N}\left(\mu_{ik}(\boldsymbol{\theta}), \sigma_{ik}^2(\boldsymbol{\theta})\right)$$

with $\mu_{ik}(\boldsymbol{\theta}) = \boldsymbol{w}_k' \boldsymbol{\beta} z_i$ and $\sigma_{ik}^2(\boldsymbol{\theta}) = \boldsymbol{w}_k' \Sigma \boldsymbol{w}_k$. Thus, the quantiles $q_{ik}^{\alpha\cdot 100\%}(\boldsymbol{\theta})$ are the quantiles of the normal distribution $\mathcal{N}\left(\mu_{ik}(\boldsymbol{\theta}), \sigma_{ik}^2(\boldsymbol{\theta})\right)$, so we can approximate the predictive quantile

$$\begin{aligned} q_{ik}^{\alpha\cdot 100\%} &= \int q_{ik}^{\alpha\cdot 100\%}(\boldsymbol{\theta}) \mathrm{d}\boldsymbol{\theta} \\ &\approx \frac{1}{5000} \sum_{r=1}^{5000} q_{ik}^{\alpha\cdot 100\%}(\boldsymbol{\theta}^{(r)}) =: \hat{q}_{ik}^{\alpha\cdot 100\%} \end{aligned}$$

using our sample from the MCMC algorithm.

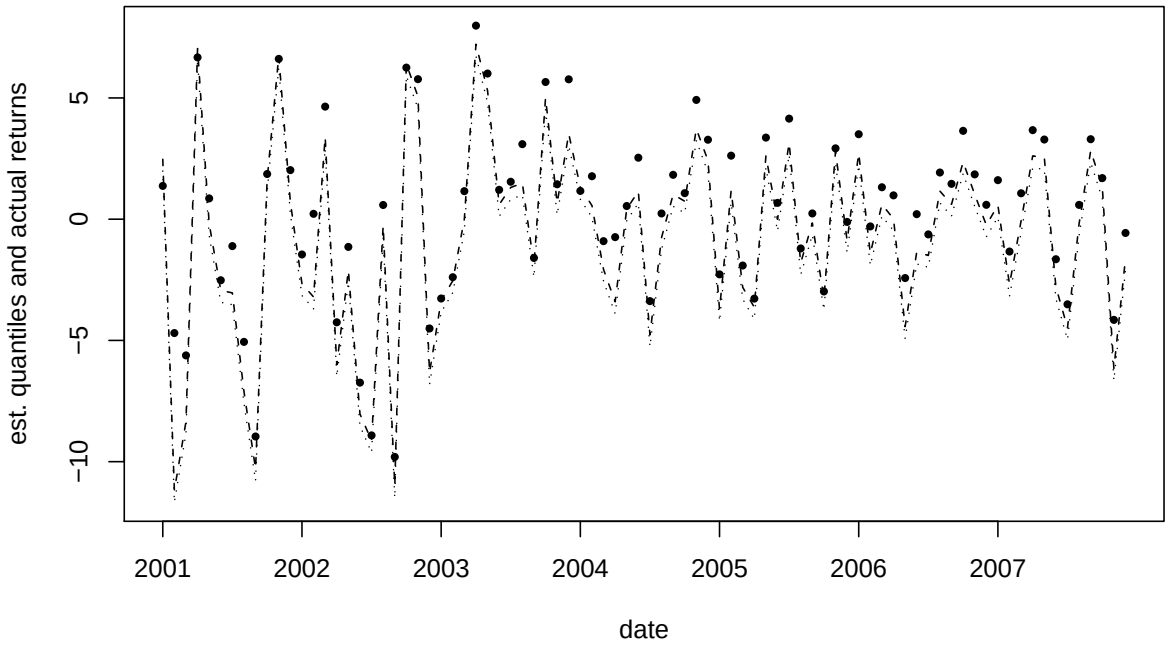

Figure 5.15: Actual returns and quantile borders (2.5% and 10%) predicted by the reduced model for the years 2001 to 2007 for the Statistic S_{i1} with equal weights

For each point in time, the estimated predictive quantile $\hat{q}_{ik}^{\alpha \cdot 100\%}$ is compared to the actual value $s_{ik} = \sum_{j=1}^{12} w_{kj} y_{ij}$ for our statistic S_{ik}. If our model has poor prediction power given market returns, it is possible that for many dates $i \in \{895, \ldots, 978\}$, the actual values are much lower than predicted by the model, i. e. smaller than q_{ik} for a small α.

The results for the statistic with equal weights, S_{i1}, are shown in Figure 5.15. The dashed line represents the estimated values of the quantile $q_{ik}^{10\%}$ and the dotted line those of $q_{ik}^{2.5\%}$ for each $i \in \{895, \ldots, 978\}$. The actual values s_{i1} of the statistic S_{i1} are shown as points. As one can see both quantile lines are close to each other, and most of the actual values lie above the lines. Over all time indices $i \in \{895, \ldots, 978\}$ there is in fact one point below the 2.5% quantile line and 5 points lie below the 10% line. As the total number of time points in the considered prediction period is 84, the expected number of points below the quantile lines is 2.1 for the 2.5% level and 8.4 for the 10% level. Thus, the results for our model based on the S_{i1} statistic are acceptable.

We do the same things for the full model of the 2nd order. The result is presented in Figure 5.16, again for the first statistic S_{i1}. We see that this graph looks very similar to that of the reduced model. Actually, the same point lies below the 2.5% quantile line, and also the five points below the 10% quantile line are the same. Thus, the satisfactory results of the reduced model do not change when all missing parameters are included.

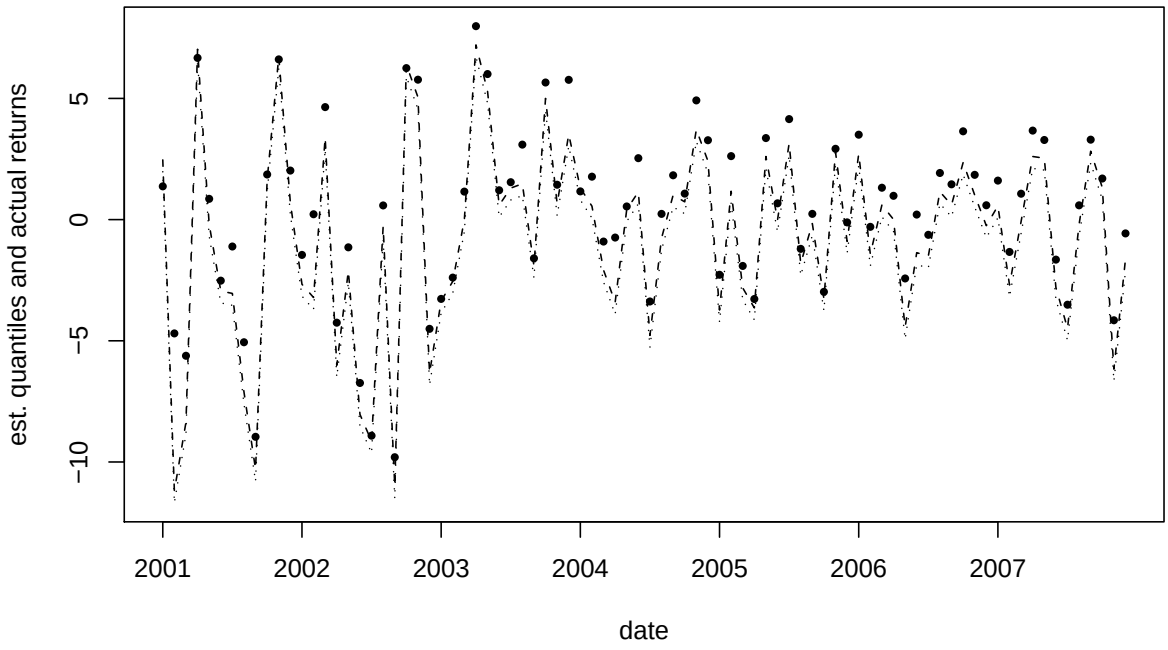

Figure 5.16: Actual returns and quantile borders (2.5% and 10%) predicted by the full model for the years 2001 to 2007 for the Statistic S_{i1} with equal weights

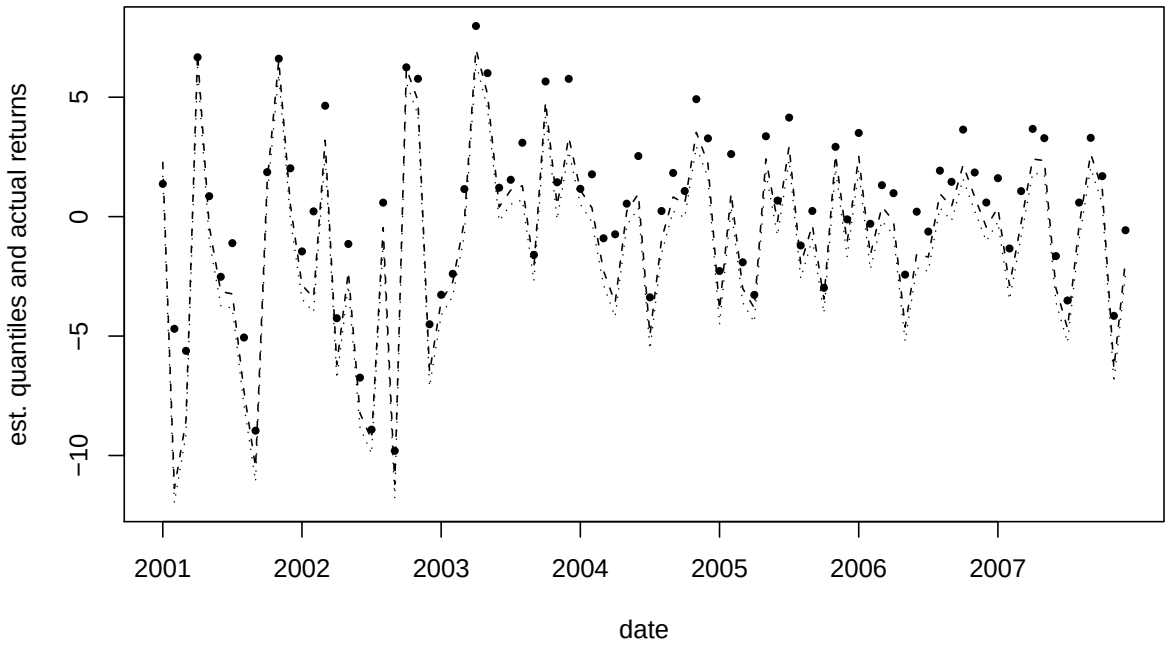

Figure 5.17: Actual returns and quantile borders (2.5% and 10%) predicted by the independence model for the years 2001 to 2007 for the Statistic S_{i1} with equal weights

In our models, we have up to 66 parameters to take into account the dependency between the residuals of the model. So it may be interesting to compare the result that we just received to a model where independence between the residuals is assumed. This independence model has only 24 parameters compared to 71 parameters in our reduced model of the second order. To get a sample for $\boldsymbol{\beta}$ and $\boldsymbol{\sigma}^2$ in the independence model, we just apply our MCMC algorithm with all components of $\boldsymbol{\rho}_{\mathcal{V}}$ fixed to 0. After that, we repeat the procedure that we developed to test the predictability of our model.

The result for the statistic with equal weights, S_{i1}, is shown in Figure 5.17. Again, the dashed line represents the estimated values of the quantile $q_{ik}^{10\%}$ and the dotted line those of $q_{ik}^{2.5\%}$ for each $i \in \{895, \ldots, 978\}$. One can see that now less points lie below the 10% quantile line. In fact, there are now only 2, compared to 5 in the models where the correlation was included. Since the expected number of actual values s_{1k} below that line is 8.4, this indicates that prediction power gets worse when the partial correlation parameters are removed from the model and independence of the portfolio return residuals is assumed. However, at least for the 2.5% quantile results stay the same, i. e. we have one point below the dotted line.

With the help of Figures 5.15 to 5.17, we analyzed the performance of the reduced, the full and the independence model concerning statistic S_{i1} representatively for all statistics $\{S_{i1}, \ldots, S_{i5}\}$. Now we want to compare the predictability of the three models over all statistics, and look in addition to $\hat{q}_{ik}^{2.5\%}$ and $\hat{q}_{ik}^{10\%}$ at the estimated quantiles with levels 1% and 5%. The results are provided in Table 5.19.

For the first statistic S_{i1}, we look at the two quantiles we have not considered yet. For both the 1% and the 5% level, the observed number of values below the quantiles are a bit better for the reduced and full model as for the independence model.

The next results we look at belong to the statistic S_{i2}, whose weights are chosen in dependence of the correlation. In exception of the 1% level where both values are equal, the observed number of values below the quantile lines of the reduced model is closer to the expected one than those of the independence model. When we compare the results of the full and the reduced model, we see only a difference for $\alpha = 2.5\%$, where the value for the reduced model is a bit better. For all other levels – and for all other statistics – we observe identical results for the full and the reduced model. So one can follow that prediction power of these two models is practically identical. This is why we compare in the following only the reduced and the independence model.

For statistic S_{i3}, which is based on positive correlations, we observe better results for the reduced model concerning the levels 1%, 2.5% and 5%. However, there are more values outside the 1% and 2.5% quantiles than for statistics S_{i1} and S_{i2}. For the 10% level, the observed number of values smaller than the estimated quantiles of the independence model are closer to the expected one in comparison to the reduced model.

Statistic	$\alpha \cdot 100\%$	Red. Model observed	Full Model observed	Ind. Model observed	expected
S_{i1}	1%	1	1	0	0.84
	2.5%	1	1	1	2.10
	5%	2	2	1	4.20
	10%	5	5	2	8.40
S_{i2}	1%	1	1	1	0.84
	2.5%	2	1	1	2.10
	5%	3	3	1	4.20
	10%	5	5	3	8.40
S_{i3}	1%	3	3	4	0.84
	2.5%	3	3	4	2.10
	5%	4	4	5	4.20
	10%	5	5	10	8.40
S_{i4}	1%	3	3	4	0.84
	2.5%	4	4	4	2.10
	5%	4	4	4	4.20
	10%	6	6	9	8.40
S_{i5}	1%	1	1	1	0.84
	2.5%	2	2	2	2.10
	5%	4	4	3	4.20
	10%	5	5	4	8.40

Table 5.19: Comparison of observed and expected number of values smaller than the estimated quantiles for different levels, models and statistics

From the results for statistic S_{i4} we see no preferred model. The results of the 1% level are better for the reduced model, those of the 10% level better for the independence model and for all other levels, the results are equal. Residual variances are taken into account when we look at statistic S_{i5}. Here we observe better values for the reduced model.

For all statistics, we observe an acceptable number of actual values below the quantiles of different levels predicted by the reduced model. This can be seen when we compare those numbers to the expected number denoted in the last column of Table 5.19. Over all statistics and levels, the reduced model shows better results than the independence model, so our validation suggests that it is worth to include the partial correlations in the model. The difference of full and reduced model is hardly observable, so the partial correlations fixed to zero in our reduced model do not improve the results and therefore may be omitted due to our validation.

Chapter 6

Conclusion and outlook

We developed an MCMC algorithm to for a joint Bayesian estimation of the regression parameters, residual variances and the correlation matrix in a multivariate regression model with correlated, normally distributed errors.

Using the literature about the pair-copula and vine concept, we modeled the dependence by a Gauss pair-copula construction arranged on a D-vine. In that context we developed an algorithm that allows to "switch" between the partial correlation specification and the correlation matrix, and proposed two ways of calculating the acceptance probability in the update of the correlation matrix: On the one hand the calculation of a D-vine likelihood, on the other hand the transformation of the partial correlation specification to the correlation matrix and afterwards the calculation of the normal likelihood. Both methods lead to similar results.

To test the algorithm, we performed an extensive small sample analysis, especially for dimensions 2 and 3. We observed convergence and well behavior of the algorithm in all considered parameter constellations and relatively accurate estimates. Furthermore, we analyzed consequences of the change of different true parameter values.

When we applied the algorithm on a data set consisting of monthly returns of 12 U.S. industrial portfolios, we also found convergence of the algorithm. We tried three different orders for the D-vine and found out that the MCMC estimates of all 3 orders and the previously determined empirical estimates of the marginal parameters and the correlation matrix were very similar.

However, the three different orders led to different possibilities of model reduction. We used the DIC criterion to compare both reduced and full models and received a preferred model. At last, we validated the model using different statistics that can be interpreted as weighted portfolios of the 12 industries and analyzed the predictive power of our model for the last 7 years of our dataset. The model showed satisfactory results and outperformed a model that assumes independence of the residuals.

There are many different ways to extend our model: At first, one could consider models with 2 or more covariates, which was not of interest in our work. Another possibility is to change the marginal distributions of the model. Instead of the normal distribution, one could use a t-distribution or a GARCH model for each marginal distribution. If only the marginal distributions are modified, one can keep the update for the partial correlations that we presented. Furthermore, the three vine orders we proposed here will lead to different models, even without model reduction. At last, one can also include other classes of pair-copulas, for example bivariate t- or Gumbel-copulas when modeling the dependence.

Appendix A

Notations

notation	definition / meaning
X, $\boldsymbol{X}$	X random variable, $\boldsymbol{X} = (X_1, \ldots, X_d)'$ random vector
x, $\boldsymbol{x}$	x realization of X, $\boldsymbol{x} = (x_1, \ldots, x_d)'$ realization of $\boldsymbol{X}$
$\boldsymbol{X}_i$	$\boldsymbol{X}_i = (X_{i1}, \ldots, X_{id})'$ random vector with time index i
$\boldsymbol{X}$	composite vector $\boldsymbol{X} = (\boldsymbol{X}_1', \ldots, \boldsymbol{X}_d')'$ or matrix $\boldsymbol{X} = (\boldsymbol{X}_1, \ldots, \boldsymbol{X}_d)$ of random vectors $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_d$
$\boldsymbol{X}_{-j}$	random vector $\boldsymbol{X}_{-j} = (X_1, \ldots, X_{j-1}, X_{j+1}, \ldots, X_d)'$
$F(x), F^{-1}(x)$	F distribution function of X, F^{-1} its inverse / quantile function
$F(x_1\|x_2)$	conditional distribution function of X_1 given X_2
$f(x_1, \ldots, x_d)$	joint density of $X_1, \ldots, X_d$
$f_j(x_j)$	marginal density of X_j
$f(x_1\|x_2)$	conditional density of X_1 given X_2
$j{:}k$	$j{:}k = j, \ldots, k$ for $j \leq k$, $j{:}k = \emptyset$ for $j > k$
$M_{-(j,k)}$	matrix M reduced by jth row and kth column
$diag(M)$	diagonal elements of matrix M
$diag(x_1, \ldots, x_d)$	diagonal matrix with diagonal elements $x_1, \ldots, x_d$
$C(\cdot,\cdot)$, $c(\cdot,\cdot)$	$C(\cdot,\cdot)$ copula, $c(\cdot,\cdot)$ copula density
$c(\cdot,\cdot\|\boldsymbol{\theta})$	copula density with copula parameter (vector) $\boldsymbol{\theta}$
$h(\cdot,\cdot,\boldsymbol{\theta})$	h-function, see (2.16)
$\Phi(\cdot)$, $\varphi(\cdot)$	$\Phi(\cdot)$ distribution function, $\varphi(\cdot)$ density of a standard normal distribution
$\Phi_{\mu,\sigma^2}(\cdot)$, $\varphi_{\mu,\sigma^2}(\cdot)$	$\Phi_{\mu,\sigma^2}(\cdot)$, distribution function, $\phi_{\mu,\sigma^2}(\cdot)$ density of $\mathcal{N}(\mu, \sigma^2)$
$\Phi_{2,\rho}(\cdot,\cdot)$, $\varphi_{2,\rho}(\cdot,\cdot)$	distribution function and density of a bivariate standard normal distribution with correlation ρ
$\Gamma(\cdot)$	$\Gamma(x) = \int_0^\infty t^{x-1} \exp\{-t\} \mathrm{d}t$ Gamma function
$B(a,b)$	$B(a,b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$ Beta function
ρ_{jk}	$\rho_{jk} = \rho(X_j, X_k)$ (product moment) correlation of X_j and X_k
ρ^{jk}	element of inverse correlation matrix in jth row and kth column
$\rho_{12;3,\ldots,d}$	partial correlation of X_1 and X_2 given $X_3, \ldots, X_d$
$\rho_{12\|3,\ldots,d}$	conditional correlation of X_1 and X_2 given $X_3, \ldots, X_d$
$\mathcal{V}$, $\mathcal{CV}$	$\mathcal{V}$ vine, $\mathcal{CV}$ constraint set, i.e. set of all edges of $\mathcal{V}$
$\boldsymbol{\rho}_{\mathcal{V}}$	set / vector of partial correlations adapted to vine $\mathcal{V}$

θ, $\boldsymbol{\theta}$	model parameter (vector) of interest
$p(\theta)$	prior density of θ
$f(\boldsymbol{x}\|\theta)$	likelihood of θ / conditional density of $\boldsymbol{x}$ given θ
$p(\theta\|\boldsymbol{x})$	posterior density of θ
$D(\theta)$	$D(\theta) = -2\log\left(f(\boldsymbol{x}\|\theta)\right)$ deviance of θ
DIC	$DIC = 2E\left(D\left(\theta\right)\|\boldsymbol{x}\right) - D\left(E(\theta\|\boldsymbol{x})\right)$ deviance information criterion
θ_{true}	true value of θ
$\hat{\theta}$	estimator for θ
$\hat{\theta}_{\alpha\cdot 100\%}$	estimated posterior $\alpha \cdot 100\%$ quantile of θ
$\hat{\theta}_{med}$, $\hat{\theta}_{mod}$	$\hat{\theta}_{med}$ estimated posterior median, $\hat{\theta}_{mod}$ of estimated posterior mode of θ
$\bar{\theta}$	estimated posterior mean of θ or sample mean of estimates $\hat{\theta}^{(1)}, \ldots, \hat{\theta}^{(r)}$
$\hat{\theta}_{IFM}$	inference for margins estimator for θ
$b(\hat{\theta})$, $\hat{b}(\hat{\theta})$	$b(\hat{\theta})$ bias, $\hat{b}(\hat{\theta})$ estimated bias of $\hat{\theta}$
$rb(\hat{\theta})$, $\widehat{rb}(\hat{\theta})$	$rb(\hat{\theta})$ relative bias, $\widehat{rb}(\hat{\theta})$ estimated relative bias of $\hat{\theta}$
$s^2(\hat{\theta})$	sample variance of estimates $\hat{\theta}^{(1)}, \ldots, \hat{\theta}^{(r)}$
$s(\bar{\theta}) = s_b(\hat{\theta})$	estimated standard error of $\hat{\theta}$ and $\hat{b}(\hat{\theta})$
$s_{rb}(\widehat{\theta_{mod}})$	estimated standard error of $\widehat{rb}(\hat{\theta})$
$\bar{p}_{acc}$	mean acceptance rate
Y_{ij}	$Y_{ij} = z_i\beta_j + \sigma_j\varepsilon_{ij}$, jth component of response variable at time i
z_i	given covariate data at time i
β_j	regression coefficient of jth component
σ_j^2	parameter for residual variance of jth component
ε_{ij}	marginally standardized residual at time i
ρ, R	ρ residual correlation parameter ($d = 2$), R residual correlation matrix
s_j^2	prior variance of β_j
a_j, b_j	prior parameters of σ_j^2
ϕ_j^2	$\phi_j^2 = \frac{1}{\sigma_j^2}$
V_{ij}	$V_{ij} = Y_{ij} - z_i\beta_j$
I_d	identity matrix of dimension d
X_i	$X_i = z_i I_d$ design matrix
D	$D = diag(\sigma_1, \ldots, \sigma_d)$
Σ	$\Sigma = DRD$ correlation matrix of $\boldsymbol{Y}_i$
$PSNR_j$	fraction of elements in $\{Y_{1j}, \ldots, Y_{nj}\}$ with $\frac{\lvert E(Y_{ij})\rvert}{\sqrt{Var(Y_{ij})}} > 2$
S_{zz}	$S_{zz} = \sum_{i=1}^n z_i^2$
S_{zy_j}	$S_{zy_j} = \sum_{i=1}^n z_i y_{ij}$
$S_{z\boldsymbol{y}}$	$S_{z\boldsymbol{y}} = \sum_{i=1}^n z_i \boldsymbol{y}_i$
$S_{\boldsymbol{v},jk}$, $S_{\boldsymbol{v}}$	$S_{\boldsymbol{v},jk} = \sum_{i=1}^n v_{ij}v_{ik}$, $S_{\boldsymbol{v}} = (S_{\boldsymbol{v},jk})_{j,k=1,\ldots,d}$
$S_{\boldsymbol{\varepsilon},jk}$	$S_{\boldsymbol{\varepsilon},jk} = \sum_{i=1}^n \varepsilon_{ij}\varepsilon_{ik}$
r_j^2	$r_j^2 = \sum_{i=1}^n (y_{ij} - \beta_j z_i)^2$

Table A.1: Notations and abbreviations

Appendix B

Tables

B.1 Results of the small sample analysis of the MCMC algorithm for 3 dimensions

Tables B.1 to B.6 show characteristics of estimates created by the MCMC algorithm using simulated data with different true parameter values and data sizes. For each scenario, the maximum value of the relative bias $\widehat{rb}(\widehat{\theta}_{mod})$ is printed in bold type. In contrast to the considered constellations in the bivariate case, we choose the same $\boldsymbol{\beta}$ configuration over all scenarios and do not consider the case where the $PSNR_j$ values are different, i. e. $PSNR_1 = PSNR_2 = PSNR_3$ holds in all scenarios. On the other hand, we also look at small data sizes, namely $n = 200$.

Sc. #	$PSNR_1$	$PSNR_2$	$PSNR_3$	n	θ	θ_{true}	$\bar{\theta}_{mod}$	$10^2 \cdot \widehat{b}(\widehat{\theta}_{mod})$	$10^2 \cdot s(\bar{\theta}_{mod})$	$10^2 \cdot \widehat{rb}(\widehat{\theta}_{mod})$	$10^2 \cdot s_{rb}(\widehat{\theta}_{mod})$
1	0.5	0.5	0.5	200	β_1	1.0000	0.9904	-0.959	1.021	-0.959	1.021
					β_2	2.0000	2.0065	0.654	1.676	0.327	0.838
					β_3	3.0000	2.9812	-1.881	3.533	-0.627	1.178
					σ_1^2	0.0625	0.0583	-0.425	0.166	-6.793	2.656
					σ_2^2	0.2500	0.2548	0.476	0.932	1.903	3.728
					σ_2^3	0.5625	0.5477	-1.480	1.964	-2.631	3.491
					ρ_{12}	0.3000	0.2861	-1.388	2.494	-4.627	8.313
					ρ_{23}	0.3000	0.2688	-3.121	2.104	**-10.404**	7.015
					$\rho_{13;2}$	0.3000	0.2983	-0.167	1.467	-0.558	4.889
2	0.5	0.5	0.5	1000	β_1	1.0000	0.9964	-0.362	0.556	-0.362	0.556
					β_2	2.0000	1.9895	-1.049	0.975	-0.525	0.487
					β_3	3.0000	3.0000	0.000	1.130	0.000	0.377
					σ_1^2	0.0625	0.0634	0.095	0.062	1.514	0.990
					σ_2^2	0.2500	0.2492	-0.078	0.397	-0.312	1.588
					σ_2^3	0.5625	0.5519	-1.058	0.508	-1.881	0.903
					ρ_{12}	0.3000	0.2879	-1.207	0.744	**-4.022**	2.480
					ρ_{23}	0.3000	0.2964	-0.361	1.043	-1.204	3.476
					$\rho_{13;2}$	0.3000	0.2923	-0.767	1.040	-2.557	3.468
3	0.5	0.5	0.5	5000	β_1	1.0000	0.9963	-0.370	0.227	-0.370	0.227
					β_2	2.0000	1.9936	-0.641	0.307	-0.321	0.154
					β_3	3.0000	2.9893	-1.072	0.511	-0.357	0.170
					σ_1^2	0.0625	0.0629	0.035	0.042	0.562	0.677
					σ_2^2	0.2500	0.2484	-0.159	0.142	-0.637	0.569
					σ_2^3	0.5625	0.5643	0.178	0.297	0.316	0.529
					ρ_{12}	0.3000	0.3062	0.618	0.346	2.060	1.153
					ρ_{23}	0.3000	0.3039	0.392	0.467	1.305	1.556
					$\rho_{13;2}$	0.3000	0.2902	-0.978	0.419	**-3.260**	1.397
4	0.5	0.5	0.5	200	β_1	1.0000	1.0130	1.301	0.814	1.301	0.814
					β_2	2.0000	2.0160	1.598	1.643	0.799	0.822
					β_3	3.0000	3.0557	5.571	3.014	1.857	1.005
					σ_1^2	0.0625	0.0607	-0.176	0.172	-2.815	2.757
					σ_2^2	0.2500	0.2365	-1.354	0.867	**-5.415**	3.469
					σ_2^3	0.5625	0.5725	1.003	2.534	1.783	4.505
					ρ_{12}	0.8000	0.7914	-0.856	1.081	-1.070	1.351
					ρ_{23}	0.3000	0.2919	-0.806	1.986	-2.687	6.620
					$\rho_{13;2}$	0.3000	0.2838	-1.617	2.195	-5.390	7.316
5	0.5	0.5	0.5	1000	β_1	1.0000	1.0037	0.368	0.449	0.368	0.449
					β_2	2.0000	2.0153	1.533	0.816	0.767	0.408
					β_3	3.0000	3.0050	0.505	1.083	0.168	0.361
					σ_1^2	0.0625	0.0637	0.116	0.073	1.854	1.174
					σ_2^2	0.2500	0.2571	0.708	0.301	2.830	1.205
					σ_2^3	0.5625	0.5789	1.639	0.652	**2.914**	1.160
					ρ_{12}	0.8000	0.8039	0.388	0.292	0.484	0.365
					ρ_{23}	0.3000	0.2934	-0.656	1.008	-2.187	3.360
					$\rho_{13;2}$	0.3000	0.3001	0.007	1.016	0.023	3.386
6	0.5	0.5	0.5	5000	β_1	1.0000	1.0005	0.050	0.216	0.050	0.216
					β_2	2.0000	1.9995	-0.054	0.266	-0.027	0.133
					β_3	3.0000	3.0067	0.671	0.751	0.224	0.250
					σ_1^2	0.0625	0.0618	-0.071	0.034	-1.140	0.538
					σ_2^2	0.2500	0.2488	-0.119	0.168	-0.477	0.672
					σ_2^3	0.5625	0.5566	-0.586	0.250	-1.042	0.445
					ρ_{12}	0.8000	0.7977	-0.230	0.148	-0.288	0.185
					ρ_{23}	0.3000	0.2995	-0.047	0.281	-0.156	0.935
					$\rho_{13;2}$	0.3000	0.2922	-0.785	0.417	**-2.616**	1.390

Table B.1: Mean of estimates, estimated bias, relative bias and their estimated standard errors for different data size and parameter constellations (Scenarios $1-6$)

Sc. #	$PSNR_1$	$PSNR_2$	$PSNR_3$	n	θ	θ_{true}	$\bar{\theta}_{mod}$	$10^2 \cdot \widehat{b}(\widehat{\theta}_{mod})$	$10^2 \cdot s(\bar{\theta}_{mod})$	$10^2 \cdot \widehat{rb}(\widehat{\theta}_{mod})$	$10^2 \cdot s_{rb}(\widehat{\theta}_{mod})$
7	0.5	0.5	0.5	200	β_1	1.0000	0.9748	-2.524	1.247	-2.524	1.247
					β_2	2.0000	1.9520	-4.799	2.007	-2.400	1.003
					β_3	3.0000	2.9792	-2.077	3.223	-0.692	1.074
					σ_1^2	0.0625	0.0608	-0.167	0.232	-2.672	3.719
					σ_2^2	0.2500	0.2376	-1.239	0.841	**-4.955**	3.365
					σ_2^3	0.5625	0.5483	-1.418	1.203	-2.521	2.139
					ρ_{12}	0.8000	0.7917	-0.835	0.868	-1.044	1.085
					ρ_{23}	0.8000	0.7818	-1.824	0.980	-2.280	1.225
					$\rho_{13;2}$	0.3000	0.3037	0.366	2.078	1.221	6.925
8	0.5	0.5	0.5	1000	β_1	1.0000	0.9974	-0.262	0.283	-0.262	0.283
					β_2	2.0000	1.9937	-0.631	0.901	-0.316	0.450
					β_3	3.0000	2.9895	-1.046	1.594	-0.349	0.531
					σ_1^2	0.0625	0.0625	0.000	0.069	0.007	1.111
					σ_2^2	0.2500	0.2439	-0.610	0.340	**-2.440**	1.358
					σ_2^3	0.5625	0.5581	-0.436	0.492	-0.776	0.874
					ρ_{12}	0.8000	0.8002	0.025	0.322	0.031	0.402
					ρ_{23}	0.8000	0.7992	-0.076	0.396	-0.095	0.495
					$\rho_{13;2}$	0.3000	0.2979	-0.206	0.794	-0.686	2.648
9	0.5	0.5	0.5	5000	β_1	1.0000	0.9991	-0.086	0.126	-0.086	0.126
					β_2	2.0000	2.0001	0.008	0.340	0.004	0.170
					β_3	3.0000	2.9947	-0.527	0.369	-0.176	0.123
					σ_1^2	0.0625	0.0620	-0.047	0.044	-0.758	0.704
					σ_2^2	0.2500	0.2489	-0.115	0.112	-0.458	0.449
					σ_2^3	0.5625	0.5606	-0.186	0.351	-0.331	0.625
					ρ_{12}	0.8000	0.7966	-0.335	0.151	-0.419	0.189
					ρ_{23}	0.8000	0.7997	-0.030	0.175	-0.037	0.218
					$\rho_{13;2}$	0.3000	0.2931	-0.694	0.378	**-2.314**	1.261
10	0.5	0.5	0.5	200	β_1	1.0000	0.9931	-0.688	1.174	-0.688	1.174
					β_2	2.0000	1.9977	-0.234	1.786	-0.117	0.893
					β_3	3.0000	3.0125	1.248	2.889	0.416	0.963
					σ_1^2	0.0625	0.0584	-0.406	0.197	-6.504	3.159
					σ_2^2	0.2500	0.2393	-1.067	0.546	-4.268	2.184
					σ_2^3	0.5625	0.5089	-5.356	1.793	**-9.522**	3.188
					ρ_{12}	0.3000	0.2962	-0.378	2.177	-1.260	7.257
					ρ_{23}	0.3000	0.3015	0.148	1.999	0.492	6.664
					$\rho_{13;2}$	0.8000	0.7871	-1.291	0.630	-1.614	0.787
11	0.5	0.5	0.5	1000	β_1	1.0000	0.9953	-0.473	0.483	-0.473	0.483
					β_2	2.0000	1.9885	-1.145	1.090	-0.573	0.545
					β_3	3.0000	2.9882	-1.178	1.190	-0.393	0.397
					σ_1^2	0.0625	0.0614	-0.109	0.053	-1.743	0.843
					σ_2^2	0.2500	0.2504	0.037	0.399	0.148	1.594
					σ_2^3	0.5625	0.5516	-1.090	0.505	**-1.939**	0.897
					ρ_{12}	0.3000	0.3047	0.472	0.868	1.573	2.895
					ρ_{23}	0.3000	0.3054	0.543	1.038	1.810	3.459
					$\rho_{13;2}$	0.8000	0.7999	-0.006	0.338	-0.007	0.423
12	0.5	0.5	0.5	5000	β_1	1.0000	1.0006	0.064	0.238	0.064	0.238
					β_2	2.0000	1.9959	-0.409	0.404	-0.204	0.202
					β_3	3.0000	3.0012	0.124	0.727	0.041	0.242
					σ_1^2	0.0625	0.0622	-0.032	0.039	-0.511	0.619
					σ_2^2	0.2500	0.2504	0.044	0.149	0.174	0.595
					σ_2^3	0.5625	0.5587	-0.377	0.361	-0.671	0.642
					ρ_{12}	0.3000	0.2997	-0.030	0.482	-0.099	1.605
					ρ_{23}	0.3000	0.3047	0.474	0.529	**1.581**	1.763
					$\rho_{13;2}$	0.8000	0.7993	-0.074	0.086	-0.093	0.108

Table B.2: Mean of estimates, estimated bias, relative bias and their estimated standard errors for different data size and parameter constellations (Scenarios $7 - 12$)

Sc. #	$PSNR_1$	$PSNR_2$	$PSNR_3$	n	θ	θ_{true}	$\bar{\theta}_{mod}$	$10^2 \cdot \widehat{b}(\widehat{\theta}_{mod})$	$10^2 \cdot s(\bar{\theta}_{mod})$	$10^2 \cdot \widehat{rb}(\widehat{\theta}_{mod})$	$10^2 \cdot s_{rb}(\widehat{\theta}_{mod})$
13	0.5	0.5	0.5	200	β_1	1.0000	0.9944	-0.561	0.839	-0.561	0.839
					β_2	2.0000	1.9998	-0.016	1.320	-0.008	0.660
					β_3	3.0000	3.0109	1.091	2.191	0.364	0.730
					σ_1^2	0.0625	0.0599	-0.264	0.147	-4.228	2.358
					σ_2^2	0.2500	0.2436	-0.641	0.309	-2.563	1.235
					σ_2^3	0.5625	0.5586	-0.389	2.305	-0.692	4.097
					ρ_{12}	0.8000	0.7871	-1.292	0.846	-1.615	1.058
					ρ_{23}	0.3000	0.2819	-1.811	1.284	**-6.036**	4.280
					$\rho_{13;2}$	0.8000	0.7934	-0.663	0.979	-0.829	1.223
14	0.5	0.5	0.5	1000	β_1	1.0000	1.0044	0.441	0.452	0.441	0.452
					β_2	2.0000	2.0088	0.881	0.981	0.440	0.490
					β_3	3.0000	3.0304	3.039	0.718	1.013	0.239
					σ_1^2	0.0625	0.0607	-0.184	0.088	-2.949	1.416
					σ_2^2	0.2500	0.2472	-0.281	0.305	-1.122	1.220
					σ_2^3	0.5625	0.5419	-2.060	0.778	-3.663	1.383
					ρ_{12}	0.8000	0.7984	-0.164	0.223	-0.205	0.278
					ρ_{23}	0.3000	0.2878	-1.217	0.876	**-4.055**	2.920
					$\rho_{13;2}$	0.8000	0.7979	-0.210	0.245	-0.263	0.306
15	0.5	0.5	0.5	5000	β_1	1.0000	1.0000	0.003	0.289	0.003	0.289
					β_2	2.0000	2.0025	0.253	0.483	0.126	0.241
					β_3	3.0000	2.9978	-0.220	0.884	-0.073	0.295
					σ_1^2	0.0625	0.0631	0.057	0.041	**0.906**	0.661
					σ_2^2	0.2500	0.2522	0.215	0.198	0.861	0.793
					σ_2^3	0.5625	0.5619	-0.063	0.325	-0.111	0.578
					ρ_{12}	0.8000	0.8018	0.180	0.172	0.225	0.215
					ρ_{23}	0.3000	0.3020	0.195	0.383	0.652	1.277
					$\rho_{13;2}$	0.8000	0.8004	0.042	0.076	0.052	0.095
16	0.5	0.5	0.5	200	β_1	1.0000	1.0027	0.269	1.046	0.269	1.046
					β_2	2.0000	1.9987	-0.129	2.306	-0.064	1.153
					β_3	3.0000	3.0097	0.972	3.652	0.324	1.217
					σ_1^2	0.0625	0.0612	-0.129	0.219	-2.070	3.506
					σ_2^2	0.2500	0.2396	-1.040	0.619	**-4.161**	2.477
					σ_2^3	0.5625	0.5664	0.390	1.757	0.693	3.124
					ρ_{12}	0.8000	0.7872	-1.283	0.684	-1.603	0.855
					ρ_{23}	0.8000	0.7858	-1.418	0.851	-1.773	1.064
					$\rho_{13;2}$	0.8000	0.8161	1.611	0.522	2.014	0.653
17	0.5	0.5	0.5	1000	β_1	1.0000	1.0064	0.645	0.562	0.645	0.562
					β_2	2.0000	2.0165	1.653	0.998	0.826	0.499
					β_3	3.0000	3.0211	2.110	1.379	0.703	0.460
					σ_1^2	0.0625	0.0616	-0.095	0.104	**-1.515**	1.661
					σ_2^2	0.2500	0.2482	-0.182	0.418	-0.728	1.673
					σ_2^3	0.5625	0.5575	-0.500	0.874	-0.889	1.554
					ρ_{12}	0.8000	0.7958	-0.419	0.289	-0.523	0.362
					ρ_{23}	0.8000	0.8001	0.015	0.230	0.018	0.287
					$\rho_{13;2}$	0.8000	0.8018	0.179	0.378	0.224	0.473
18	0.5	0.5	0.5	5000	β_1	1.0000	0.9974	-0.264	0.177	-0.264	0.177
					β_2	2.0000	1.9950	-0.495	0.355	-0.248	0.178
					β_3	3.0000	2.9894	-1.056	0.532	**-0.352**	0.177
					σ_1^2	0.0625	0.0625	-0.003	0.043	-0.048	0.687
					σ_2^2	0.2500	0.2497	-0.030	0.163	-0.119	0.650
					σ_2^3	0.5625	0.5612	-0.134	0.378	-0.238	0.671
					ρ_{12}	0.8000	0.7974	-0.261	0.129	-0.326	0.162
					ρ_{23}	0.8000	0.7990	-0.105	0.117	-0.131	0.146
					$\rho_{13;2}$	0.8000	0.8011	0.106	0.148	0.133	0.186

Table B.3: Mean of estimates, estimated bias, relative bias and their estimated standard errors for different data size and parameter constellations (Scenarios $13 - 18$)

Sc. #	$PSNR_1$	$PSNR_2$	$PSNR_3$	n	θ	θ_{true}	$\bar{\theta}_{mod}$	$10^2 \cdot \widehat{b}(\widehat{\theta}_{mod})$	$10^2 \cdot s(\bar{\theta}_{mod})$	$10^2 \cdot \widehat{rb}(\widehat{\theta}_{mod})$	$10^2 \cdot s_{rb}(\widehat{\theta}_{mod})$
19	0.8	0.8	0.8	200	β_1	1.00	0.9985	-0.149	0.321	-0.149	0.321
					β_2	2.00	2.0044	0.436	0.850	0.218	0.425
					β_3	3.00	2.9864	-1.355	1.312	-0.452	0.437
					σ_1^2	0.01	0.0102	0.016	0.016	1.615	1.626
					σ_2^2	0.04	0.0416	0.164	0.127	4.101	3.174
					σ_2^3	0.09	0.0932	0.323	0.157	3.593	1.749
					ρ_{12}	0.30	0.3280	2.801	2.698	**9.338**	8.994
					ρ_{23}	0.30	0.3044	0.442	1.669	1.474	5.564
					$\rho_{13;2}$	0.30	0.3021	0.207	2.433	0.690	8.109
20	0.8	0.8	0.8	1000	β_1	1.00	0.9999	-0.014	0.225	-0.014	0.225
					β_2	2.00	1.9980	-0.201	0.355	-0.101	0.178
					β_3	3.00	3.0003	0.026	0.757	0.009	0.252
					σ_1^2	0.01	0.0100	0.000	0.008	-0.034	0.825
					σ_2^2	0.04	0.0398	-0.023	0.048	-0.575	1.212
					σ_2^3	0.09	0.0885	-0.148	0.133	-1.641	1.482
					ρ_{12}	0.30	0.3213	2.129	0.510	**7.098**	1.700
					ρ_{23}	0.30	0.2969	-0.308	1.265	-1.027	4.217
					$\rho_{13;2}$	0.30	0.3084	0.837	1.131	2.790	3.769
21	0.8	0.8	0.8	5000	β_1	1.00	1.0002	0.018	0.071	0.018	0.071
					β_2	2.00	1.9985	-0.149	0.165	-0.075	0.083
					β_3	3.00	3.0012	0.122	0.136	0.041	0.045
					σ_1^2	0.01	0.0101	0.005	0.006	0.511	0.618
					σ_2^2	0.04	0.0398	-0.024	0.029	-0.592	0.723
					σ_2^3	0.09	0.0896	-0.045	0.040	-0.499	0.442
					ρ_{12}	0.30	0.2946	-0.540	0.372	**-1.801**	1.242
					ρ_{23}	0.30	0.2985	-0.147	0.399	-0.490	1.330
					$\rho_{13;2}$	0.30	0.2953	-0.465	0.494	-1.551	1.648
22	0.8	0.8	0.8	200	β_1	1.00	1.0090	0.899	0.473	0.899	0.473
					β_2	2.00	2.0086	0.861	0.882	0.431	0.441
					β_3	3.00	3.0221	2.213	1.209	0.738	0.403
					σ_1^2	0.01	0.0096	-0.042	0.033	-4.159	3.274
					σ_2^2	0.04	0.0386	-0.145	0.067	-3.621	1.677
					σ_2^3	0.09	0.0877	-0.227	0.344	-2.519	3.820
					ρ_{12}	0.80	0.7859	-1.407	0.389	-1.759	0.487
					ρ_{23}	0.30	0.2502	-4.978	1.714	**-16.594**	5.715
					$\rho_{13;2}$	0.30	0.2805	-1.946	2.287	-6.488	7.623
23	0.8	0.8	0.8	1000	β_1	1.00	1.0026	0.260	0.163	0.260	0.163
					β_2	2.00	2.0039	0.388	0.386	0.194	0.193
					β_3	3.00	3.0080	0.801	0.560	0.267	0.187
					σ_1^2	0.01	0.0100	0.003	0.011	0.253	1.093
					σ_2^2	0.04	0.0396	-0.037	0.041	-0.937	1.015
					σ_2^3	0.09	0.0887	-0.130	0.101	-1.442	1.126
					ρ_{12}	0.80	0.7948	-0.518	0.282	-0.648	0.352
					ρ_{23}	0.30	0.2899	-1.006	1.207	-3.354	4.025
					$\rho_{13;2}$	0.30	0.3146	1.456	1.188	**4.852**	3.960
24	0.8	0.8	0.8	5000	β_1	1.00	0.9999	-0.008	0.059	-0.008	0.059
					β_2	2.00	2.0001	0.013	0.158	0.006	0.079
					β_3	3.00	3.0010	0.097	0.256	0.032	0.085
					σ_1^2	0.01	0.0099	-0.009	0.004	-0.918	0.416
					σ_2^2	0.04	0.0397	-0.031	0.027	-0.772	0.686
					σ_2^3	0.09	0.0895	-0.049	0.038	-0.548	0.418
					ρ_{12}	0.80	0.7985	-0.149	0.148	-0.186	0.186
					ρ_{23}	0.30	0.2960	-0.396	0.269	**-1.319**	0.896
					$\rho_{13;2}$	0.30	0.2990	-0.103	0.594	-0.344	1.980

Table B.4: Mean of estimates, estimated bias, relative bias and their estimated standard errors for different data size and parameter constellations (Scenarios $19 - 24$)

Sc. #	$PSNR_1$	$PSNR_2$	$PSNR_3$	n	θ	θ_{true}	$\bar{\theta}_{mod}$	$10^2 \cdot \widehat{b}(\widehat{\theta}_{mod})$	$10^2 \cdot s(\bar{\theta}_{mod})$	$10^2 \cdot \widehat{rb}(\widehat{\theta}_{mod})$	$10^2 \cdot s_{rb}(\widehat{\theta}_{mod})$
25	0.8	0.8	0.8	200	β_1	1.00	1.0009	0.093	0.324	0.093	0.324
					β_2	2.00	2.0022	0.224	0.528	0.112	0.264
					β_3	3.00	2.9921	-0.794	0.828	-0.265	0.276
					σ_1^2	0.01	0.0095	-0.047	0.044	-4.722	4.415
					σ_2^2	0.04	0.0381	-0.186	0.189	-4.641	4.736
					σ_2^3	0.09	0.0835	-0.646	0.323	**-7.180**	3.593
					ρ_{12}	0.80	0.7768	-2.323	1.274	-2.904	1.593
					ρ_{23}	0.80	0.7960	-0.404	0.741	-0.505	0.926
					$\rho_{13;2}$	0.30	0.2799	-2.012	2.333	-6.708	7.775
26	0.8	0.8	0.8	1000	β_1	1.00	0.9996	-0.041	0.097	-0.041	0.097
					β_2	2.00	2.0019	0.193	0.221	0.096	0.111
					β_3	3.00	2.9996	-0.038	0.284	-0.013	0.095
					σ_1^2	0.01	0.0101	0.008	0.016	0.780	1.550
					σ_2^2	0.04	0.0403	0.030	0.053	0.758	1.328
					σ_2^3	0.09	0.0890	-0.104	0.116	-1.151	1.292
					ρ_{12}	0.80	0.8003	0.029	0.354	0.036	0.443
					ρ_{23}	0.80	0.8018	0.178	0.286	0.222	0.358
					$\rho_{13;2}$	0.30	0.3088	0.876	1.147	**2.921**	3.824
27	0.8	0.8	0.8	5000	β_1	1.00	0.9995	-0.049	0.075	-0.049	0.075
					β_2	2.00	1.9984	-0.161	0.119	-0.080	0.060
					β_3	3.00	2.9973	-0.267	0.206	-0.089	0.069
					σ_1^2	0.01	0.0101	0.008	0.008	0.792	0.767
					σ_2^2	0.04	0.0403	0.032	0.025	0.792	0.633
					σ_2^3	0.09	0.0907	0.067	0.069	0.748	0.770
					ρ_{12}	0.80	0.8022	0.221	0.151	0.277	0.188
					ρ_{23}	0.80	0.8013	0.130	0.142	0.162	0.178
					$\rho_{13;2}$	0.30	0.3080	0.804	0.290	**2.679**	0.967
28	0.8	0.8	0.8	200	β_1	1.00	0.9983	-0.166	0.452	-0.166	0.452
					β_2	2.00	2.0000	-0.004	0.681	-0.002	0.340
					β_3	3.00	2.9954	-0.461	1.507	-0.154	0.502
					σ_1^2	0.01	0.0100	-0.004	0.030	-0.430	2.952
					σ_2^2	0.04	0.0397	-0.034	0.106	-0.843	2.644
					σ_2^3	0.09	0.0875	-0.248	0.310	-2.758	3.442
					ρ_{12}	0.30	0.3204	2.040	1.905	6.801	6.349
					ρ_{23}	0.30	0.3217	2.166	0.992	**7.221**	3.306
					$\rho_{13;2}$	0.80	0.7931	-0.694	0.659	-0.868	0.824
29	0.8	0.8	0.8	1000	β_1	1.00	0.9998	-0.023	0.142	-0.023	0.142
					β_2	2.00	1.9973	-0.272	0.285	-0.136	0.142
					β_3	3.00	2.9925	-0.754	0.362	-0.251	0.121
					σ_1^2	0.01	0.0100	-0.001	0.016	-0.139	1.589
					σ_2^2	0.04	0.0389	-0.110	0.050	**-2.749**	1.250
					σ_2^3	0.09	0.0902	0.018	0.151	0.198	1.673
					ρ_{12}	0.30	0.3040	0.400	1.024	1.332	3.414
					ρ_{23}	0.30	0.3052	0.517	1.005	1.723	3.351
					$\rho_{13;2}$	0.80	0.7991	-0.085	0.277	-0.107	0.347
30	0.8	0.8	0.8	5000	β_1	1.00	1.0009	0.093	0.079	0.093	0.079
					β_2	2.00	1.9996	-0.036	0.099	-0.018	0.049
					β_3	3.00	3.0022	0.218	0.240	0.073	0.080
					σ_1^2	0.01	0.0099	-0.007	0.010	-0.705	0.972
					σ_2^2	0.04	0.0399	-0.007	0.019	-0.185	0.478
					σ_2^3	0.09	0.0890	-0.101	0.062	-1.118	0.685
					ρ_{12}	0.30	0.2955	-0.450	0.323	-1.499	1.077
					ρ_{23}	0.30	0.2936	-0.643	0.415	**-2.144**	1.385
					$\rho_{13;2}$	0.80	0.7971	-0.293	0.216	-0.367	0.270

Table B.5: Mean of estimates, estimated bias, relative bias and their estimated standard errors for different data size and parameter constellations (Scenarios $25 - 30$)

Sc. #	$PSNR_1$	$PSNR_2$	$PSNR_3$	n	θ	θ_{true}	$\bar{\theta}_{mod}$	$10^2 \cdot \widehat{b}(\widehat{\theta}_{mod})$	$10^2 \cdot s(\bar{\theta}_{mod})$	$10^2 \cdot \widehat{rb}(\widehat{\theta}_{mod})$	$10^2 \cdot s_{rb}(\widehat{\theta}_{mod})$
31	0.8	0.8	0.8	200	β_1	1.00	0.9966	-0.340	0.305	-0.340	0.305
					β_2	2.00	1.9940	-0.601	0.489	-0.300	0.244
					β_3	3.00	2.9927	-0.730	1.255	-0.243	0.418
					σ_1^2	0.01	0.0096	-0.044	0.043	**-4.406**	4.260
					σ_2^2	0.04	0.0388	-0.117	0.149	-2.914	3.725
					σ_2^3	0.09	0.0862	-0.375	0.278	-4.168	3.083
					ρ_{12}	0.80	0.7904	-0.957	1.311	-1.197	1.639
					ρ_{23}	0.30	0.2902	-0.984	2.632	-3.279	8.772
					$\rho_{13;2}$	0.80	0.7894	-1.063	1.103	-1.329	1.378
32	0.8	0.8	0.8	1000	β_1	1.00	0.9992	-0.083	0.176	-0.083	0.176
					β_2	2.00	1.9998	-0.017	0.483	-0.008	0.241
					β_3	3.00	2.9954	-0.457	0.307	-0.152	0.102
					σ_1^2	0.01	0.0099	-0.009	0.011	-0.884	1.068
					σ_2^2	0.04	0.0389	-0.110	0.067	**-2.743**	1.680
					σ_2^3	0.09	0.0916	0.159	0.148	1.763	1.649
					ρ_{12}	0.80	0.7963	-0.375	0.424	-0.469	0.530
					ρ_{23}	0.30	0.2996	-0.036	0.783	-0.119	2.609
					$\rho_{13;2}$	0.80	0.8031	0.311	0.354	0.389	0.442
33	0.8	0.8	0.8	5000	β_1	1.00	0.9995	-0.051	0.078	-0.051	0.078
					β_2	2.00	1.9993	-0.065	0.146	-0.033	0.073
					β_3	3.00	3.0004	0.039	0.198	0.013	0.066
					σ_1^2	0.01	0.0100	-0.001	0.004	-0.117	0.389
					σ_2^2	0.04	0.0401	0.008	0.027	0.188	0.672
					σ_2^3	0.09	0.0895	-0.050	0.037	**-0.560**	0.409
					ρ_{12}	0.80	0.8007	0.070	0.180	0.088	0.225
					ρ_{23}	0.30	0.3015	0.155	0.393	0.516	1.312
					$\rho_{13;2}$	0.80	0.7980	-0.197	0.083	-0.246	0.104
34	0.8	0.8	0.8	200	β_1	1.00	0.9961	-0.387	0.549	-0.387	0.549
					β_2	2.00	1.9967	-0.327	0.866	-0.163	0.433
					β_3	3.00	2.9880	-1.198	1.440	-0.399	0.480
					σ_1^2	0.01	0.0091	-0.088	0.036	**-8.771**	3.649
					σ_2^2	0.04	0.0368	-0.319	0.088	-7.976	2.191
					σ_2^3	0.09	0.0833	-0.674	0.298	-7.493	3.316
					ρ_{12}	0.80	0.7900	-0.997	0.889	-1.246	1.111
					ρ_{23}	0.80	0.7895	-1.054	1.126	-1.318	1.408
					$\rho_{13;2}$	0.80	0.7950	-0.500	1.182	-0.624	1.477
35	0.8	0.8	0.8	1000	β_1	1.00	0.9996	-0.043	0.112	-0.043	0.112
					β_2	2.00	1.9966	-0.341	0.238	-0.170	0.119
					β_3	3.00	2.9985	-0.150	0.445	-0.050	0.148
					σ_1^2	0.01	0.0099	-0.013	0.014	**-1.264**	1.442
					σ_2^2	0.04	0.0396	-0.040	0.051	-1.003	1.274
					σ_2^3	0.09	0.0889	-0.113	0.111	-1.255	1.232
					ρ_{12}	0.80	0.7935	-0.646	0.438	-0.808	0.547
					ρ_{23}	0.80	0.7947	-0.531	0.442	-0.664	0.552
					$\rho_{13;2}$	0.80	0.8031	0.314	0.330	0.393	0.413
36	0.8	0.8	0.8	5000	β_1	1.00	1.0001	0.007	0.053	0.007	0.053
					β_2	2.00	1.9996	-0.041	0.129	-0.020	0.064
					β_3	3.00	3.0008	0.079	0.244	0.026	0.081
					σ_1^2	0.01	0.0100	-0.004	0.006	**-0.448**	0.625
					σ_2^2	0.04	0.0400	0.004	0.020	0.098	0.502
					σ_2^3	0.09	0.0899	-0.013	0.052	-0.145	0.576
					ρ_{12}	0.80	0.7981	-0.190	0.211	-0.238	0.264
					ρ_{23}	0.80	0.7981	-0.194	0.207	-0.242	0.258
					$\rho_{13;2}$	0.80	0.8022	0.223	0.195	0.279	0.243

Table B.6: Mean of estimates, estimated bias, relative bias and their estimated standard errors for different data size and parameter constellations (Scenarios $31 - 36$)

B.2 MCMC estimates of parameters for the U.S. industrial returns data

We have defined three D-vine orders and run our MCMC algorithm for each order. The results for the marginal parameters have already been provided in Tables 5.14 to 5.16, those for the copula parameter are presented in tables B.8 to B.10. For the reduced models, the characteristics of the MCMC estimates are shown for both marginal and copula parameters in Tables B.11 to B.22.

To avoid long notations of the partial correlations with large conditioned sets, we do not use the portfolio indices in $\mathcal{I} = \{N, D, M, E, C, B, T, U, S, H, \$, O\}$ in the tables but instead the position indices $\{1, \ldots, 12\}$. Of course, this implicates that partial correlations with equal position indices in different orders usually also refer to different partial correlations.

For example, $\rho_{1,5|\cdot} = \rho_{1,5|2,3,4}$ denotes $\rho_{NC|DME}$ in the first construction, $\rho_{US|\$BE}$ in the second and $\rho_{BD|\$NE}$ in the third construction. Table B.7 shows for each position index the corresponding portfolio index in the three orders.

Another notation we use in the copula parameter tables is the quantity $\mathrm{C}_{\alpha\cdot 100\%}$ for $\alpha \cdot 100\% \in \{2.5\%, 5\%\}$. This quantity is 1 if the parameter θ in the corresponding row is credible on the level $\alpha \cdot 100\%$, meaning that $0 \notin \left[\theta_{(0.5\alpha)\cdot 100\%}, \theta_{(1-0.5\alpha)\cdot 100\%}\right]$. Otherwise, $\mathrm{C}_{\alpha\cdot 100\%}$ is 0.

Constr.	1	2	3	4	5	6	7	8	9	10	11	12
1	*N*	*D*	*M*	*E*	*C*	*B*	*T*	*U*	*S*	*H*	\$	*O*
2	*U*	\$	*B*	*E*	*S*	*N*	*H*	*D*	*C*	*O*	*M*	*T*
3	*B*	\$	*N*	*E*	*D*	*T*	*O*	*C*	*U*	*S*	*H*	*M*

Table B.7: Positions of covariables in the D-vine constructions 1, 2 and 3. For each row, the first tree of the respective D-vine can be derived by connecting adjacent columns.

θ	$\hat{\theta}_{min}$	$\hat{\theta}_{2.5\%}$	$\hat{\theta}_{5\%}$	$\hat{\theta}_{med}$	$\bar{\theta}$	$\hat{\theta}_{95\%}$	$\hat{\theta}_{97.5\%}$	$\hat{\theta}_{max}$	$\hat{\theta}_{mod}$	$\hat{\theta}_{IFM}$	$C_{10\%}$	$C_{5\%}$	$\bar{p}_{acc}$
$\rho_{1,2}$	-0.121	-0.090	-0.083	-0.032	-0.031	0.022	0.036	0.086	-0.032	-0.031	0	0	0.50
$\rho_{2,3}$	0.170	0.207	0.213	0.257	0.256	0.302	0.311	0.325	0.255	0.260	1	1	0.52
$\rho_{3,4}$	-0.145	-0.128	-0.120	-0.067	-0.065	-0.011	-0.002	0.015	-0.068	-0.066	1	1	0.51
$\rho_{4,5}$	-0.147	-0.130	-0.117	-0.062	-0.063	-0.015	-0.004	0.023	-0.059	-0.066	1	1	0.49
$\rho_{5,6}$	-0.220	-0.182	-0.171	-0.118	-0.119	-0.070	-0.059	-0.026	-0.118	-0.123	1	1	0.51
$\rho_{6,7}$	-0.153	-0.127	-0.111	-0.062	-0.062	-0.015	-0.005	0.036	-0.061	-0.060	1	1	0.53
$\rho_{7,8}$	0.039	0.071	0.079	0.130	0.130	0.189	0.198	0.228	0.125	0.132	1	1	0.56
$\rho_{8,9}$	-0.153	-0.113	-0.106	-0.057	-0.057	-0.011	0.002	0.025	-0.055	-0.055	1	0	0.50
$\rho_{9,10}$	-0.005	0.055	0.071	0.114	0.114	0.163	0.178	0.203	0.114	0.114	1	1	0.53
$\rho_{10,11}$	0.008	0.026	0.037	0.089	0.088	0.140	0.148	0.172	0.089	0.090	1	1	0.54
$\rho_{11,12}$	0.023	0.049	0.056	0.101	0.102	0.154	0.172	0.191	0.097	0.103	1	1	0.53
$\rho_{1,3\vert\cdot}$	-0.093	-0.068	-0.059	-0.003	-0.002	0.055	0.063	0.099	-0.004	-0.001	0	0	0.50
$\rho_{2,4\vert\cdot}$	-0.249	-0.218	-0.212	-0.164	-0.162	-0.111	-0.099	-0.058	-0.162	-0.162	1	1	0.55
$\rho_{3,5\vert\cdot}$	0.136	0.178	0.192	0.240	0.239	0.286	0.291	0.319	0.242	0.243	1	1	0.49
$\rho_{4,6\vert\cdot}$	-0.408	-0.390	-0.380	-0.339	-0.338	-0.288	-0.283	-0.245	-0.342	-0.339	1	1	0.46
$\rho_{5,7\vert\cdot}$	-0.207	-0.171	-0.166	-0.120	-0.118	-0.068	-0.064	-0.034	-0.122	-0.125	1	1	0.53
$\rho_{6,8\vert\cdot}$	-0.289	-0.264	-0.258	-0.212	-0.211	-0.158	-0.149	-0.132	-0.214	-0.216	1	1	0.54
$\rho_{7,9\vert\cdot}$	-0.071	-0.046	-0.032	0.024	0.023	0.076	0.083	0.102	0.027	0.020	0	0	0.52
$\rho_{8,10\vert\cdot}$	-0.072	-0.032	-0.016	0.039	0.038	0.088	0.095	0.116	0.038	0.038	0	0	0.53
$\rho_{9,11\vert\cdot}$	0.047	0.060	0.076	0.129	0.127	0.178	0.187	0.227	0.129	0.130	1	1	0.53
$\rho_{10,12\vert\cdot}$	-0.223	-0.193	-0.188	-0.136	-0.135	-0.080	-0.072	-0.042	-0.136	-0.131	1	1	0.53
$\rho_{1,4\vert\cdot}$	-0.275	-0.249	-0.236	-0.187	-0.187	-0.134	-0.124	-0.093	-0.187	-0.195	1	1	0.49
$\rho_{2,5\vert\cdot}$	0.058	0.071	0.082	0.134	0.134	0.183	0.200	0.255	0.136	0.136	1	1	0.55
$\rho_{3,6\vert\cdot}$	-0.089	-0.035	-0.023	0.024	0.024	0.076	0.085	0.110	0.024	0.024	0	0	0.49
$\rho_{4,7\vert\cdot}$	-0.331	-0.312	-0.301	-0.256	-0.256	-0.206	-0.189	-0.171	-0.256	-0.263	1	1	0.48
$\rho_{5,8\vert\cdot}$	-0.116	-0.077	-0.071	-0.019	-0.019	0.029	0.039	0.061	-0.019	-0.022	0	0	0.55
$\rho_{6,9\vert\cdot}$	-0.181	-0.154	-0.147	-0.093	-0.093	-0.041	-0.031	0.013	-0.094	-0.094	1	1	0.51
$\rho_{7,10\vert\cdot}$	-0.128	-0.085	-0.079	-0.027	-0.024	0.038	0.048	0.068	-0.032	-0.028	0	0	0.55
$\rho_{8,11\vert\cdot}$	0.291	0.311	0.322	0.372	0.370	0.415	0.422	0.444	0.373	0.378	1	1	0.50
$\rho_{9,12\vert\cdot}$	-0.144	-0.108	-0.100	-0.051	-0.052	0.000	0.010	0.074	-0.050	-0.052	0	0	0.53
$\rho_{1,5\vert\cdot}$	0.105	0.131	0.143	0.189	0.189	0.238	0.250	0.285	0.188	0.184	1	1	0.50
$\rho_{2,6\vert\cdot}$	-0.137	-0.092	-0.081	-0.030	-0.031	0.019	0.024	0.061	-0.030	-0.033	0	0	0.52
$\rho_{3,7\vert\cdot}$	-0.406	-0.375	-0.369	-0.324	-0.324	-0.281	-0.267	-0.244	-0.321	-0.329	1	1	0.49
$\rho_{4,8\vert\cdot}$	-0.061	-0.039	-0.031	0.020	0.020	0.068	0.077	0.104	0.021	0.018	0	0	0.52
$\rho_{5,9\vert\cdot}$	-0.089	-0.042	-0.032	0.019	0.019	0.071	0.083	0.091	0.019	0.018	0	0	0.52
$\rho_{6,10\vert\cdot}$	-0.125	-0.083	-0.075	-0.025	-0.024	0.029	0.038	0.066	-0.025	-0.027	0	0	0.54
$\rho_{7,11\vert\cdot}$	-0.069	-0.056	-0.047	0.006	0.006	0.059	0.075	0.088	0.006	0.004	0	0	0.57
$\rho_{8,12\vert\cdot}$	-0.202	-0.184	-0.174	-0.121	-0.121	-0.068	-0.062	-0.005	-0.121	-0.118	1	1	0.54
$\rho_{1,6\vert\cdot}$	-0.481	-0.439	-0.433	-0.391	-0.389	-0.341	-0.332	-0.301	-0.392	-0.400	1	1	0.47
$\rho_{2,7\vert\cdot}$	-0.178	-0.140	-0.134	-0.080	-0.080	-0.029	-0.020	-0.005	-0.079	-0.083	1	1	0.56
$\rho_{3,8\vert\cdot}$	-0.247	-0.227	-0.220	-0.168	-0.168	-0.112	-0.105	-0.070	-0.171	-0.170	1	1	0.54
$\rho_{4,9\vert\cdot}$	-0.436	-0.419	-0.417	-0.378	-0.376	-0.331	-0.322	-0.301	-0.378	-0.383	1	1	0.46
$\rho_{5,10\vert\cdot}$	0.041	0.079	0.089	0.139	0.138	0.185	0.196	0.227	0.138	0.135	1	1	0.54
$\rho_{6,11\vert\cdot}$	-0.225	-0.202	-0.190	-0.140	-0.142	-0.095	-0.087	-0.043	-0.138	-0.142	1	1	0.56
$\rho_{7,12\vert\cdot}$	-0.218	-0.173	-0.162	-0.116	-0.115	-0.063	-0.058	-0.027	-0.116	-0.113	1	1	0.53
$\rho_{1,7\vert\cdot}$	-0.191	-0.167	-0.155	-0.107	-0.107	-0.060	-0.042	-0.025	-0.110	-0.116	1	1	0.54
$\rho_{2,8\vert\cdot}$	-0.117	-0.080	-0.072	-0.024	-0.024	0.027	0.037	0.064	-0.023	-0.026	0	0	0.56
$\rho_{3,9\vert\cdot}$	-0.223	-0.191	-0.181	-0.125	-0.126	-0.074	-0.063	-0.030	-0.124	-0.126	1	1	0.54
$\rho_{4,10\vert\cdot}$	-0.198	-0.179	-0.169	-0.120	-0.116	-0.059	-0.048	-0.017	-0.123	-0.127	1	1	0.54
$\rho_{5,11\vert\cdot}$	-0.136	-0.119	-0.106	-0.063	-0.061	-0.011	-0.003	0.024	-0.062	-0.061	1	1	0.56
$\rho_{6,12\vert\cdot}$	-0.227	-0.205	-0.198	-0.144	-0.143	-0.092	-0.081	-0.044	-0.144	-0.141	1	1	0.54
$\rho_{1,8\vert\cdot}$	-0.034	-0.009	0.006	0.050	0.051	0.098	0.110	0.145	0.048	0.049	1	0	0.54
$\rho_{2,9\vert\cdot}$	0.004	0.043	0.049	0.101	0.102	0.151	0.162	0.190	0.103	0.103	1	1	0.54
$\rho_{3,10\vert\cdot}$	-0.293	-0.279	-0.270	-0.222	-0.221	-0.175	-0.166	-0.128	-0.222	-0.226	1	1	0.53
$\rho_{4,11\vert\cdot}$	-0.269	-0.226	-0.219	-0.159	-0.161	-0.109	-0.098	-0.068	-0.157	-0.164	1	1	0.55
$\rho_{5,12\vert\cdot}$	-0.247	-0.225	-0.218	-0.164	-0.164	-0.118	-0.107	-0.061	-0.161	-0.165	1	1	0.54
$\rho_{1,9\vert\cdot}$	0.267	0.290	0.301	0.349	0.348	0.394	0.398	0.417	0.349	0.347	1	1	0.50
$\rho_{2,10\vert\cdot}$	-0.298	-0.253	-0.245	-0.198	-0.196	-0.145	-0.134	-0.097	-0.199	-0.200	1	1	0.54
$\rho_{3,11\vert\cdot}$	-0.051	-0.012	0.001	0.054	0.053	0.102	0.110	0.136	0.054	0.052	1	0	0.56
$\rho_{4,12\vert\cdot}$	-0.328	-0.309	-0.299	-0.255	-0.254	-0.206	-0.200	-0.156	-0.256	-0.255	1	1	0.53
$\rho_{1,10\vert\cdot}$	0.171	0.206	0.213	0.264	0.263	0.310	0.318	0.344	0.264	0.262	1	1	0.53
$\rho_{2,11\vert\cdot}$	-0.074	-0.051	-0.043	0.001	0.003	0.060	0.070	0.112	-0.001	0.001	0	0	0.57
$\rho_{3,12\vert\cdot}$	0.164	0.196	0.211	0.258	0.259	0.309	0.316	0.339	0.258	0.262	1	1	0.54
$\rho_{1,11\vert\cdot}$	-0.043	-0.008	-0.001	0.049	0.051	0.107	0.120	0.135	0.049	0.050	0	0	0.58
$\rho_{2,12\vert\cdot}$	-0.224	-0.205	-0.194	-0.149	-0.148	-0.097	-0.084	-0.058	-0.150	-0.149	1	1	0.57
$\rho_{1,12\vert\cdot}$	-0.107	-0.084	-0.078	-0.020	-0.019	0.036	0.050	0.080	-0.020	-0.020	0	0	0.57

Table B.8: MCMC results for the copula parameters of the 1st vine construction.

θ	$\hat{\theta}_{min}$	$\hat{\theta}_{2.5\%}$	$\hat{\theta}_{5\%}$	$\hat{\theta}_{med}$	$\bar{\theta}$	$\hat{\theta}_{95\%}$	$\hat{\theta}_{97.5\%}$	$\hat{\theta}_{max}$	$\hat{\theta}_{mod}$	$\hat{\theta}_{IFM}$	$C_{10\%}$	$C_{5\%}$	$\bar{p}_{acc}$
$\rho_{1,2}$	0.279	0.305	0.311	0.364	0.361	0.401	0.411	0.433	0.367	0.367	1	1	0.50
$\rho_{2,3}$	-0.312	-0.281	-0.272	-0.221	-0.221	-0.173	-0.163	-0.138	-0.220	-0.224	1	1	0.54
$\rho_{3,4}$	-0.409	-0.378	-0.370	-0.326	-0.325	-0.281	-0.273	-0.238	-0.325	-0.328	1	1	0.47
$\rho_{4,5}$	-0.413	-0.379	-0.367	-0.315	-0.318	-0.268	-0.260	-0.218	-0.314	-0.324	1	1	0.50
$\rho_{5,6}$	0.330	0.349	0.358	0.410	0.409	0.453	0.463	0.476	0.410	0.414	1	1	0.47
$\rho_{6,7}$	0.261	0.275	0.281	0.332	0.330	0.374	0.380	0.418	0.334	0.332	1	1	0.52
$\rho_{7,8}$	-0.241	-0.219	-0.205	-0.154	-0.154	-0.098	-0.088	-0.049	-0.156	-0.156	1	1	0.54
$\rho_{8,9}$	0.072	0.123	0.145	0.192	0.191	0.238	0.246	0.270	0.192	0.197	1	1	0.53
$\rho_{9,10}$	-0.224	-0.203	-0.195	-0.144	-0.146	-0.100	-0.090	-0.049	-0.141	-0.147	1	1	0.53
$\rho_{10,11}$	0.206	0.236	0.243	0.292	0.291	0.336	0.343	0.386	0.294	0.294	1	1	0.50
$\rho_{11,12}$	-0.387	-0.373	-0.365	-0.320	-0.320	-0.270	-0.260	-0.239	-0.321	-0.325	1	1	0.50
$\rho_{1,3\vert\cdot}$	-0.241	-0.212	-0.204	-0.153	-0.153	-0.110	-0.100	-0.043	-0.152	-0.154	1	1	0.55
$\rho_{2,4\vert\cdot}$	-0.284	-0.240	-0.236	-0.186	-0.186	-0.133	-0.122	-0.078	-0.185	-0.193	1	1	0.54
$\rho_{3,5\vert\cdot}$	-0.284	-0.256	-0.251	-0.203	-0.203	-0.153	-0.145	-0.121	-0.203	-0.208	1	1	0.51
$\rho_{4,6\vert\cdot}$	-0.157	-0.119	-0.107	-0.056	-0.056	-0.007	0.004	0.036	-0.056	-0.061	1	0	0.53
$\rho_{5,7\vert\cdot}$	-0.111	-0.089	-0.082	-0.033	-0.032	0.022	0.034	0.064	-0.036	-0.027	0	0	0.56
$\rho_{6,8\vert\cdot}$	-0.079	-0.042	-0.029	0.022	0.022	0.073	0.083	0.116	0.022	0.023	0	0	0.56
$\rho_{7,9\vert\cdot}$	0.088	0.119	0.137	0.178	0.179	0.225	0.233	0.266	0.174	0.178	1	1	0.56
$\rho_{8,10\vert\cdot}$	-0.125	-0.086	-0.074	-0.023	-0.023	0.027	0.039	0.070	-0.022	-0.019	0	0	0.55
$\rho_{9,11\vert\cdot}$	0.200	0.243	0.253	0.300	0.300	0.347	0.358	0.381	0.299	0.306	1	1	0.51
$\rho_{10,12\vert\cdot}$	-0.107	-0.081	-0.069	-0.025	-0.022	0.029	0.041	0.063	-0.026	-0.021	0	0	0.54
$\rho_{1,4\vert\cdot}$	-0.016	0.000	0.012	0.057	0.059	0.112	0.119	0.135	0.057	0.059	1	0	0.56
$\rho_{2,5\vert\cdot}$	-0.040	-0.011	0.001	0.059	0.057	0.110	0.117	0.149	0.062	0.058	1	0	0.56
$\rho_{3,6\vert\cdot}$	-0.447	-0.418	-0.409	-0.355	-0.358	-0.314	-0.307	-0.286	-0.354	-0.367	1	1	0.50
$\rho_{4,7\vert\cdot}$	-0.183	-0.152	-0.147	-0.093	-0.093	-0.042	-0.033	-0.017	-0.091	-0.097	1	1	0.55
$\rho_{5,8\vert\cdot}$	0.072	0.107	0.118	0.175	0.172	0.218	0.226	0.254	0.180	0.175	1	1	0.55
$\rho_{6,9\vert\cdot}$	0.049	0.074	0.084	0.134	0.135	0.190	0.200	0.213	0.133	0.133	1	1	0.56
$\rho_{7,10\vert\cdot}$	-0.208	-0.166	-0.159	-0.107	-0.109	-0.055	-0.043	-0.020	-0.107	-0.108	1	1	0.56
$\rho_{8,11\vert\cdot}$	0.153	0.177	0.188	0.240	0.240	0.294	0.298	0.333	0.240	0.244	1	1	0.54
$\rho_{9,12\vert\cdot}$	-0.124	-0.101	-0.094	-0.042	-0.043	0.006	0.017	0.037	-0.038	-0.046	0	0	0.55
$\rho_{1,5\vert\cdot}$	-0.212	-0.178	-0.163	-0.109	-0.109	-0.055	-0.050	-0.038	-0.109	-0.111	1	1	0.56
$\rho_{2,6\vert\cdot}$	-0.002	0.016	0.028	0.075	0.079	0.134	0.139	0.176	0.073	0.082	1	1	0.56
$\rho_{3,7\vert\cdot}$	-0.043	-0.020	-0.006	0.037	0.039	0.090	0.098	0.117	0.035	0.036	0	0	0.55
$\rho_{4,8\vert\cdot}$	-0.232	-0.208	-0.201	-0.151	-0.154	-0.104	-0.099	-0.070	-0.150	-0.159	1	1	0.54
$\rho_{5,9\vert\cdot}$	-0.180	-0.148	-0.138	-0.089	-0.089	-0.037	-0.026	0.006	-0.088	-0.089	1	1	0.57
$\rho_{6,10\vert\cdot}$	-0.034	0.003	0.014	0.065	0.065	0.123	0.130	0.149	0.065	0.067	1	1	0.56
$\rho_{7,11\vert\cdot}$	-0.222	-0.194	-0.184	-0.130	-0.130	-0.075	-0.069	-0.041	-0.130	-0.132	1	1	0.55
$\rho_{8,12\vert\cdot}$	-0.136	-0.097	-0.086	-0.040	-0.037	0.020	0.034	0.055	-0.042	-0.037	0	0	0.55
$\rho_{1,6\vert\cdot}$	-0.033	0.003	0.014	0.060	0.061	0.116	0.125	0.137	0.059	0.060	1	1	0.56
$\rho_{2,7\vert\cdot}$	-0.068	-0.028	-0.019	0.028	0.028	0.077	0.082	0.128	0.028	0.027	0	0	0.57
$\rho_{3,8\vert\cdot}$	-0.189	-0.128	-0.123	-0.072	-0.071	-0.021	-0.016	0.014	-0.072	-0.078	1	1	0.55
$\rho_{4,9\vert\cdot}$	-0.086	-0.067	-0.055	-0.006	-0.005	0.049	0.059	0.080	-0.010	-0.005	0	0	0.55
$\rho_{5,10\vert\cdot}$	-0.146	-0.122	-0.109	-0.057	-0.059	-0.012	-0.004	0.028	-0.057	-0.061	1	1	0.55
$\rho_{6,11\vert\cdot}$	-0.102	-0.070	-0.064	-0.011	-0.011	0.045	0.051	0.107	-0.011	-0.013	0	0	0.57
$\rho_{7,12\vert\cdot}$	-0.175	-0.139	-0.126	-0.072	-0.074	-0.024	-0.018	0.021	-0.072	-0.078	1	1	0.56
$\rho_{1,7\vert\cdot}$	-0.101	-0.065	-0.055	-0.011	-0.008	0.045	0.057	0.110	-0.015	-0.009	0	0	0.57
$\rho_{2,8\vert\cdot}$	-0.109	-0.072	-0.065	-0.017	-0.016	0.038	0.046	0.074	-0.018	-0.019	0	0	0.58
$\rho_{3,9\vert\cdot}$	-0.183	-0.153	-0.145	-0.085	-0.086	-0.031	-0.015	0.017	-0.086	-0.088	1	1	0.55
$\rho_{4,10\vert\cdot}$	-0.225	-0.201	-0.195	-0.142	-0.141	-0.089	-0.078	-0.053	-0.144	-0.138	1	1	0.55
$\rho_{5,11\vert\cdot}$	-0.147	-0.124	-0.116	-0.066	-0.065	-0.017	-0.008	0.022	-0.064	-0.063	1	1	0.58
$\rho_{6,12\vert\cdot}$	-0.064	-0.021	-0.013	0.043	0.042	0.096	0.106	0.136	0.045	0.043	0	0	0.55
$\rho_{1,8\vert\cdot}$	-0.144	-0.118	-0.109	-0.057	-0.058	-0.009	-0.001	0.036	-0.057	-0.059	1	1	0.57
$\rho_{2,9\vert\cdot}$	-0.169	-0.144	-0.136	-0.094	-0.092	-0.045	-0.034	-0.005	-0.096	-0.093	1	1	0.57
$\rho_{3,10\vert\cdot}$	-0.313	-0.277	-0.268	-0.214	-0.216	-0.166	-0.157	-0.118	-0.212	-0.214	1	1	0.54
$\rho_{4,11\vert\cdot}$	-0.120	-0.078	-0.069	-0.020	-0.020	0.027	0.042	0.098	-0.018	-0.020	0	0	0.55
$\rho_{5,12\vert\cdot}$	-0.089	-0.067	-0.060	-0.001	-0.003	0.047	0.056	0.080	0.003	-0.003	0	0	0.56
$\rho_{1,9\vert\cdot}$	-0.123	-0.057	-0.051	-0.005	-0.006	0.043	0.052	0.090	-0.007	-0.008	0	0	0.57
$\rho_{2,10\vert\cdot}$	-0.063	-0.005	0.002	0.052	0.052	0.103	0.112	0.132	0.052	0.053	1	0	0.56
$\rho_{3,11\vert\cdot}$	0.006	0.031	0.045	0.096	0.095	0.143	0.152	0.182	0.098	0.095	1	1	0.55
$\rho_{4,12\vert\cdot}$	-0.366	-0.343	-0.333	-0.281	-0.282	-0.231	-0.220	-0.179	-0.280	-0.292	1	1	0.53
$\rho_{1,10\vert\cdot}$	-0.251	-0.200	-0.196	-0.142	-0.143	-0.086	-0.079	-0.059	-0.140	-0.148	1	1	0.56
$\rho_{2,11\vert\cdot}$	-0.134	-0.092	-0.080	-0.032	-0.031	0.020	0.029	0.044	-0.033	-0.032	0	0	0.57
$\rho_{3,12\vert\cdot}$	-0.309	-0.286	-0.269	-0.222	-0.221	-0.172	-0.164	-0.094	-0.222	-0.228	1	1	0.55
$\rho_{1,11\vert\cdot}$	-0.272	-0.236	-0.226	-0.170	-0.169	-0.115	-0.107	-0.077	-0.171	-0.170	1	1	0.56
$\rho_{2,12\vert\cdot}$	-0.085	-0.069	-0.058	-0.004	-0.003	0.049	0.061	0.102	-0.002	-0.004	0	0	0.57
$\rho_{1,12\vert\cdot}$	-0.044	-0.018	-0.009	0.039	0.042	0.101	0.109	0.125	0.039	0.044	0	0	0.57

Table B.9: MCMC results for the copula parameters of the 2nd vine construction.

θ	$\hat{\theta}_{min}$	$\hat{\theta}_{2.5\%}$	$\hat{\theta}_{5\%}$	$\hat{\theta}_{med}$	$\bar{\theta}$	$\hat{\theta}_{95\%}$	$\hat{\theta}_{97.5\%}$	$\hat{\theta}_{max}$	$\hat{\theta}_{mod}$	$\hat{\theta}_{IFM}$	$C_{10\%}$	$C_{5\%}$	$\bar{p}_{acc}$
$\rho_{1,2}$	-0.307	-0.281	-0.268	-0.223	-0.223	-0.178	-0.171	-0.120	-0.225	-0.224	1	1	0.49
$\rho_{2,3}$	0.113	0.149	0.160	0.207	0.206	0.249	0.263	0.298	0.209	0.212	1	1	0.50
$\rho_{3,4}$	-0.258	-0.235	-0.227	-0.181	-0.178	-0.127	-0.123	-0.100	-0.184	-0.187	1	1	0.48
$\rho_{4,5}$	-0.254	-0.218	-0.215	-0.168	-0.168	-0.120	-0.112	-0.080	-0.167	-0.174	1	1	0.52
$\rho_{5,6}$	-0.229	-0.177	-0.172	-0.119	-0.119	-0.069	-0.057	-0.023	-0.116	-0.120	1	1	0.54
$\rho_{6,7}$	-0.232	-0.178	-0.173	-0.115	-0.115	-0.061	-0.046	-0.031	-0.116	-0.115	1	1	0.54
$\rho_{7,8}$	-0.232	-0.214	-0.203	-0.149	-0.150	-0.100	-0.086	-0.047	-0.148	-0.147	1	1	0.52
$\rho_{8,9}$	-0.090	-0.069	-0.057	-0.006	-0.006	0.048	0.061	0.080	-0.008	-0.008	0	0	0.52
$\rho_{9,10}$	-0.152	-0.124	-0.109	-0.054	-0.055	-0.008	0.004	0.019	-0.051	-0.055	1	0	0.52
$\rho_{10,11}$	0.032	0.048	0.059	0.113	0.112	0.161	0.174	0.208	0.114	0.114	1	1	0.54
$\rho_{11,12}$	-0.231	-0.212	-0.202	-0.151	-0.151	-0.101	-0.083	-0.043	-0.149	-0.153	1	1	0.51
$\rho_{1,3\vert\cdot}$	-0.359	-0.344	-0.332	-0.282	-0.283	-0.235	-0.220	-0.166	-0.280	-0.287	1	1	0.49
$\rho_{2,4\vert\cdot}$	-0.189	-0.122	-0.115	-0.069	-0.067	-0.015	-0.005	0.016	-0.070	-0.067	1	1	0.52
$\rho_{3,5\vert\cdot}$	-0.162	-0.131	-0.115	-0.063	-0.065	-0.017	-0.006	0.008	-0.061	-0.065	1	1	0.52
$\rho_{4,6\vert\cdot}$	-0.334	-0.288	-0.280	-0.235	-0.234	-0.183	-0.175	-0.135	-0.234	-0.238	1	1	0.52
$\rho_{5,7\vert\cdot}$	-0.140	-0.128	-0.121	-0.068	-0.066	-0.012	-0.007	0.034	-0.070	-0.062	1	1	0.54
$\rho_{6,8\vert\cdot}$	-0.213	-0.189	-0.179	-0.130	-0.130	-0.083	-0.072	-0.037	-0.131	-0.135	1	1	0.54
$\rho_{7,9\vert\cdot}$	-0.155	-0.124	-0.109	-0.063	-0.063	-0.013	0.000	0.029	-0.066	-0.066	1	0	0.52
$\rho_{8,10\vert\cdot}$	-0.072	-0.041	-0.029	0.028	0.026	0.073	0.087	0.119	0.031	0.027	0	0	0.51
$\rho_{9,11\vert\cdot}$	-0.059	-0.031	-0.016	0.038	0.037	0.086	0.094	0.133	0.037	0.038	0	0	0.53
$\rho_{10,12\vert\cdot}$	-0.126	-0.093	-0.081	-0.033	-0.032	0.019	0.026	0.071	-0.036	-0.029	0	0	0.52
$\rho_{1,4\vert\cdot}$	-0.515	-0.488	-0.468	-0.428	-0.428	-0.384	-0.373	-0.349	-0.428	-0.435	1	1	0.46
$\rho_{2,5\vert\cdot}$	-0.082	-0.055	-0.049	0.003	0.002	0.052	0.062	0.094	0.001	0.001	0	0	0.54
$\rho_{3,6\vert\cdot}$	-0.104	-0.086	-0.078	-0.031	-0.032	0.015	0.025	0.049	-0.028	-0.040	0	0	0.52
$\rho_{4,7\vert\cdot}$	-0.215	-0.188	-0.181	-0.138	-0.135	-0.081	-0.073	-0.041	-0.142	-0.133	1	1	0.53
$\rho_{5,8\vert\cdot}$	0.101	0.120	0.129	0.177	0.177	0.232	0.243	0.277	0.177	0.178	1	1	0.53
$\rho_{6,9\vert\cdot}$	0.048	0.073	0.081	0.127	0.126	0.170	0.178	0.214	0.129	0.125	1	1	0.54
$\rho_{7,10\vert\cdot}$	-0.138	-0.107	-0.099	-0.051	-0.051	-0.004	0.005	0.051	-0.054	-0.049	1	0	0.51
$\rho_{8,11\vert\cdot}$	0.022	0.069	0.084	0.143	0.140	0.191	0.202	0.231	0.146	0.140	1	1	0.54
$\rho_{9,12\vert\cdot}$	-0.274	-0.253	-0.245	-0.197	-0.196	-0.146	-0.137	-0.122	-0.199	-0.197	1	1	0.51
$\rho_{1,5\vert\cdot}$	-0.185	-0.150	-0.135	-0.085	-0.087	-0.035	-0.021	0.001	-0.083	-0.092	1	1	0.55
$\rho_{2,6\vert\cdot}$	-0.051	-0.019	-0.011	0.041	0.041	0.098	0.108	0.137	0.043	0.040	0	0	0.55
$\rho_{3,7\vert\cdot}$	-0.117	-0.094	-0.081	-0.028	-0.028	0.033	0.042	0.078	-0.029	-0.023	0	0	0.51
$\rho_{4,8\vert\cdot}$	-0.166	-0.141	-0.134	-0.077	-0.078	-0.021	-0.013	0.014	-0.072	-0.080	1	1	0.54
$\rho_{5,9\vert\cdot}$	-0.149	-0.118	-0.113	-0.064	-0.062	-0.006	0.002	0.027	-0.067	-0.064	1	0	0.53
$\rho_{6,10\vert\cdot}$	-0.079	-0.038	-0.032	0.019	0.019	0.068	0.080	0.110	0.017	0.017	0	0	0.51
$\rho_{7,11\vert\cdot}$	-0.186	-0.156	-0.144	-0.095	-0.095	-0.042	-0.037	-0.014	-0.095	-0.095	1	1	0.53
$\rho_{8,12\vert\cdot}$	0.197	0.222	0.227	0.275	0.274	0.313	0.324	0.383	0.278	0.278	1	1	0.51
$\rho_{1,6\vert\cdot}$	-0.261	-0.231	-0.220	-0.172	-0.171	-0.121	-0.115	-0.080	-0.172	-0.181	1	1	0.55
$\rho_{2,7\vert\cdot}$	-0.003	0.050	0.055	0.107	0.105	0.149	0.153	0.184	0.108	0.106	1	1	0.55
$\rho_{3,8\vert\cdot}$	0.103	0.116	0.126	0.175	0.176	0.225	0.232	0.287	0.173	0.175	1	1	0.51
$\rho_{4,9\vert\cdot}$	-0.013	0.022	0.029	0.080	0.080	0.131	0.140	0.184	0.080	0.079	1	1	0.52
$\rho_{5,10\vert\cdot}$	0.046	0.080	0.091	0.139	0.139	0.189	0.206	0.258	0.139	0.144	1	1	0.51
$\rho_{6,11\vert\cdot}$	-0.110	-0.085	-0.076	-0.027	-0.025	0.026	0.038	0.059	-0.026	-0.024	0	0	0.54
$\rho_{7,12\vert\cdot}$	0.247	0.262	0.272	0.320	0.319	0.363	0.369	0.399	0.321	0.326	1	1	0.51
$\rho_{1,7\vert\cdot}$	-0.298	-0.277	-0.268	-0.212	-0.214	-0.164	-0.158	-0.124	-0.211	-0.214	1	1	0.55
$\rho_{2,8\vert\cdot}$	-0.138	-0.110	-0.100	-0.054	-0.051	0.002	0.016	0.027	-0.057	-0.052	0	0	0.56
$\rho_{3,9\vert\cdot}$	0.042	0.062	0.074	0.131	0.129	0.179	0.187	0.224	0.135	0.135	1	1	0.51
$\rho_{4,10\vert\cdot}$	-0.386	-0.363	-0.360	-0.311	-0.311	-0.264	-0.259	-0.232	-0.309	-0.315	1	1	0.49
$\rho_{5,11\vert\cdot}$	-0.285	-0.267	-0.259	-0.215	-0.211	-0.159	-0.147	-0.122	-0.217	-0.214	1	1	0.53
$\rho_{6,12\vert\cdot}$	-0.380	-0.326	-0.316	-0.268	-0.268	-0.224	-0.205	-0.189	-0.270	-0.272	1	1	0.52
$\rho_{1,8\vert\cdot}$	-0.232	-0.205	-0.200	-0.153	-0.152	-0.101	-0.088	-0.053	-0.153	-0.157	1	1	0.56
$\rho_{2,9\vert\cdot}$	0.284	0.308	0.315	0.360	0.361	0.406	0.417	0.432	0.359	0.366	1	1	0.51
$\rho_{3,10\vert\cdot}$	0.313	0.352	0.358	0.401	0.403	0.451	0.458	0.469	0.399	0.407	1	1	0.48
$\rho_{4,11\vert\cdot}$	-0.241	-0.221	-0.210	-0.162	-0.161	-0.108	-0.102	-0.079	-0.165	-0.168	1	1	0.55
$\rho_{5,12\vert\cdot}$	0.102	0.134	0.147	0.194	0.195	0.246	0.252	0.270	0.192	0.197	1	1	0.56
$\rho_{1,9\vert\cdot}$	-0.195	-0.162	-0.155	-0.106	-0.108	-0.058	-0.049	-0.026	-0.106	-0.110	1	1	0.57
$\rho_{2,10\vert\cdot}$	-0.001	0.025	0.035	0.085	0.086	0.140	0.149	0.179	0.085	0.085	1	1	0.58
$\rho_{3,11\vert\cdot}$	0.190	0.229	0.237	0.279	0.278	0.319	0.330	0.353	0.280	0.278	1	1	0.53
$\rho_{4,12\vert\cdot}$	-0.179	-0.151	-0.143	-0.090	-0.090	-0.045	-0.028	-0.003	-0.087	-0.093	1	1	0.57
$\rho_{1,10\vert\cdot}$	-0.211	-0.165	-0.152	-0.100	-0.099	-0.048	-0.035	0.013	-0.098	-0.102	1	1	0.57
$\rho_{2,11\vert\cdot}$	-0.041	-0.014	-0.004	0.051	0.050	0.100	0.112	0.145	0.054	0.052	0	0	0.58
$\rho_{3,12\vert\cdot}$	-0.043	-0.014	0.000	0.049	0.050	0.100	0.111	0.138	0.051	0.047	0	0	0.57
$\rho_{1,11\vert\cdot}$	-0.123	-0.069	-0.063	-0.009	-0.010	0.046	0.054	0.081	-0.010	-0.015	0	0	0.57
$\rho_{2,12\vert\cdot}$	-0.072	-0.038	-0.028	0.022	0.023	0.074	0.082	0.127	0.023	0.024	0	0	0.56
$\rho_{1,12\vert\cdot}$	-0.078	-0.055	-0.047	0.001	0.001	0.053	0.062	0.086	-0.002	0.000	0	0	0.57

Table B.10: MCMC results for the copula parameters of the 3rd vine construction.

θ	$\hat{\theta}_{min}$	$\hat{\theta}_{2.5\%}$	$\hat{\theta}_{5\%}$	$\hat{\theta}_{med}$	$\bar{\theta}$	$\hat{\theta}_{95\%}$	$\hat{\theta}_{97.5\%}$	$\hat{\theta}_{max}$	$\hat{\theta}_{mod}$	$\hat{\theta}_{IFM}$	$\bar{p}_{acc}$
$\beta_1(=\beta_N)$	0.726	0.742	0.745	0.767	0.767	0.789	0.795	0.800	0.767	0.767	1.000
$\beta_2(=\beta_D)$	1.150	1.174	1.181	1.221	1.221	1.258	1.267	1.291	1.223	1.218	1.000
$\beta_3(=\beta_M)$	1.158	1.167	1.170	1.190	1.190	1.211	1.215	1.227	1.190	1.189	1.000
$\beta_4(=\beta_E)$	0.798	0.817	0.824	0.858	0.858	0.895	0.900	0.914	0.859	0.858	1.000
$\beta_5(=\beta_C)$	0.933	0.948	0.956	0.978	0.978	1.003	1.010	1.020	0.977	0.978	1.000
$\beta_6(=\beta_B)$	1.238	1.254	1.258	1.295	1.294	1.329	1.336	1.347	1.296	1.294	1.000
$\beta_7(=\beta_T)$	0.599	0.608	0.613	0.643	0.643	0.674	0.678	0.697	0.643	0.643	1.000
$\beta_8(=\beta_U)$	0.736	0.763	0.769	0.806	0.807	0.843	0.851	0.875	0.806	0.805	1.000
$\beta_9(=\beta_S)$	0.910	0.928	0.935	0.960	0.960	0.986	0.992	1.005	0.958	0.961	1.000
$\beta_{10}(=\beta_H)$	0.808	0.826	0.831	0.864	0.865	0.901	0.910	0.929	0.864	0.865	1.000
$\beta_{11}(=\beta_\$)$	1.110	1.125	1.130	1.161	1.159	1.184	1.191	1.204	1.164	1.158	1.000
$\beta_{12}(=\beta_O)$	1.099	1.106	1.108	1.135	1.135	1.163	1.165	1.184	1.135	1.135	1.000
$\sigma^2_1(=\sigma^2_N)$	4.397	4.584	4.641	4.978	4.998	5.329	5.420	5.843	4.972	5.017	0.280
$\sigma^2_2(=\sigma^2_D)$	12.686	13.115	13.301	14.477	14.467	15.635	15.995	16.955	14.491	14.427	0.299
$\sigma^2_3(=\sigma^2_M)$	3.556	3.770	3.812	4.080	4.084	4.385	4.442	4.618	4.079	4.083	0.291
$\sigma^2_4(=\sigma^2_E)$	12.699	13.327	13.599	14.595	14.576	15.589	15.863	16.374	14.587	14.412	0.274
$\sigma^2_5(=\sigma^2_C)$	5.819	6.056	6.122	6.587	6.600	7.084	7.186	7.605	6.580	6.548	0.294
$\sigma^2_6(=\sigma^2_B)$	9.501	9.691	9.808	10.502	10.538	11.409	11.580	11.849	10.457	10.521	0.282
$\sigma^2_7(=\sigma^2_T)$	7.917	8.277	8.392	8.990	8.996	9.634	9.706	9.912	8.981	8.942	0.298
$\sigma^2_8(=\sigma^2_U)$	11.955	12.381	12.513	13.516	13.526	14.736	14.938	15.364	13.545	13.456	0.294
$\sigma^2_9(=\sigma^2_S)$	6.929	7.206	7.317	7.860	7.883	8.490	8.568	9.467	7.857	7.829	0.290
$\sigma^2_{10}(=\sigma^2_H)$	10.346	10.893	11.048	11.799	11.810	12.703	12.800	13.209	11.782	11.803	0.303
$\sigma^2_{11}(=\sigma^2_\$)$	6.746	6.987	7.044	7.548	7.577	8.170	8.322	8.713	7.523	7.527	0.294
$\sigma^2_{12}(=\sigma^2_O)$	5.984	6.264	6.370	6.833	6.848	7.339	7.443	7.828	6.785	6.790	0.297

Table B.11: MCMC results for the marginal parameters of the reduced model of the 1st vine construction with credible level 10%.

θ	$\hat{\theta}_{min}$	$\hat{\theta}_{2.5\%}$	$\hat{\theta}_{5\%}$	$\hat{\theta}_{med}$	$\bar{\theta}$	$\hat{\theta}_{95\%}$	$\hat{\theta}_{97.5\%}$	$\hat{\theta}_{max}$	$\hat{\theta}_{mod}$	$\hat{\theta}_{IFM}$	$\bar{p}_{acc}$
$\beta_1(=\beta_N)$	0.728	0.741	0.747	0.766	0.766	0.787	0.792	0.804	0.765	0.767	1.000
$\beta_2(=\beta_D)$	1.150	1.170	1.180	1.215	1.216	1.253	1.258	1.279	1.215	1.218	1.000
$\beta_3(=\beta_M)$	1.160	1.166	1.170	1.189	1.188	1.209	1.212	1.227	1.190	1.189	1.000
$\beta_4(=\beta_E)$	0.792	0.816	0.826	0.860	0.861	0.894	0.900	0.928	0.860	0.858	1.000
$\beta_5(=\beta_C)$	0.944	0.949	0.955	0.978	0.978	1.001	1.006	1.014	0.978	0.978	1.000
$\beta_6(=\beta_B)$	1.243	1.257	1.261	1.291	1.291	1.323	1.327	1.342	1.292	1.294	1.000
$\beta_7(=\beta_T)$	0.600	0.610	0.615	0.645	0.645	0.676	0.679	0.693	0.647	0.643	1.000
$\beta_8(=\beta_U)$	0.735	0.767	0.777	0.808	0.808	0.844	0.851	0.872	0.808	0.805	1.000
$\beta_9(=\beta_S)$	0.910	0.927	0.932	0.961	0.960	0.989	0.994	1.003	0.963	0.961	1.000
$\beta_{10}(=\beta_H)$	0.811	0.822	0.831	0.865	0.866	0.900	0.905	0.929	0.864	0.865	1.000
$\beta_{11}(=\beta_\$)$	1.119	1.127	1.131	1.160	1.160	1.189	1.192	1.209	1.159	1.158	1.000
$\beta_{12}(=\beta_O)$	1.086	1.104	1.108	1.135	1.134	1.158	1.161	1.176	1.136	1.135	1.000
$\sigma^2_1(=\sigma^2_N)$	4.440	4.601	4.683	4.981	5.009	5.386	5.487	5.663	4.960	5.017	0.285
$\sigma^2_2(=\sigma^2_D)$	12.629	13.302	13.403	14.424	14.438	15.552	15.755	16.740	14.412	14.427	0.307
$\sigma^2_3(=\sigma^2_M)$	3.623	3.782	3.818	4.088	4.108	4.435	4.478	4.662	4.066	4.083	0.285
$\sigma^2_4(=\sigma^2_E)$	13.035	13.298	13.515	14.541	14.534	15.580	15.775	16.557	14.527	14.412	0.273
$\sigma^2_5(=\sigma^2_C)$	5.848	6.074	6.131	6.594	6.601	7.138	7.238	7.718	6.581	6.548	0.296
$\sigma^2_6(=\sigma^2_B)$	9.299	9.639	9.743	10.556	10.569	11.470	11.730	11.961	10.559	10.521	0.283
$\sigma^2_7(=\sigma^2_T)$	7.964	8.254	8.368	8.935	8.968	9.633	9.779	10.271	8.905	8.942	0.292
$\sigma^2_8(=\sigma^2_U)$	11.682	12.309	12.519	13.448	13.470	14.430	14.664	15.739	13.478	13.456	0.294
$\sigma^2_9(=\sigma^2_S)$	7.149	7.250	7.330	7.898	7.900	8.465	8.609	8.957	7.892	7.829	0.288
$\sigma^2_{10}(=\sigma^2_H)$	10.699	10.937	11.057	11.778	11.803	12.703	12.881	13.445	11.745	11.803	0.293
$\sigma^2_{11}(=\sigma^2_\$)$	6.358	6.914	7.017	7.560	7.574	8.170	8.289	8.775	7.534	7.527	0.303
$\sigma^2_{12}(=\sigma^2_O)$	6.035	6.306	6.336	6.842	6.847	7.337	7.410	7.977	6.852	6.790	0.298

Table B.12: MCMC results for the marginal parameters of the reduced model of the 1st vine construction with credible level 5%.

θ	$\hat{\theta}_{min}$	$\hat{\theta}_{2.5\%}$	$\hat{\theta}_{5\%}$	$\hat{\theta}_{med}$	$\bar{\theta}$	$\hat{\theta}_{95\%}$	$\hat{\theta}_{97.5\%}$	$\hat{\theta}_{max}$	$\hat{\theta}_{mod}$	$\hat{\theta}_{IFM}$	$C_{10\%}$	$C_{5\%}$	$\bar{p}_{acc}$
$\rho_{2,3}$	0.175	0.197	0.206	0.256	0.255	0.300	0.310	0.331	0.256	0.260	1	1	0.53
$\rho_{3,4}$	-0.143	-0.122	-0.111	-0.061	-0.060	-0.012	0.001	0.039	-0.061	-0.066	1	0	0.51
$\rho_{4,5}$	-0.158	-0.126	-0.115	-0.060	-0.061	-0.007	0.004	0.030	-0.059	-0.066	1	0	0.49
$\rho_{5,6}$	-0.200	-0.176	-0.166	-0.121	-0.120	-0.071	-0.062	-0.032	-0.120	-0.123	1	1	0.50
$\rho_{6,7}$	-0.129	-0.110	-0.100	-0.057	-0.057	-0.006	0.003	0.031	-0.058	-0.060	1	0	0.53
$\rho_{7,8}$	0.026	0.072	0.080	0.128	0.129	0.182	0.189	0.218	0.127	0.132	1	1	0.55
$\rho_{8,9}$	-0.142	-0.104	-0.100	-0.051	-0.052	-0.005	0.002	0.069	-0.051	-0.055	1	0	0.50
$\rho_{9,10}$	0.003	0.051	0.059	0.113	0.114	0.162	0.170	0.196	0.111	0.114	1	1	0.54
$\rho_{10,11}$	-0.007	0.016	0.025	0.070	0.071	0.118	0.122	0.144	0.070	0.090	1	1	0.54
$\rho_{11,12}$	0.007	0.048	0.054	0.103	0.102	0.148	0.159	0.205	0.101	0.103	1	1	0.52
$\rho_{2,4\vert\cdot}$	-0.266	-0.223	-0.218	-0.165	-0.167	-0.122	-0.110	-0.064	-0.163	-0.162	1	1	0.54
$\rho_{3,5\vert\cdot}$	0.142	0.175	0.186	0.238	0.236	0.287	0.292	0.306	0.238	0.243	1	1	0.49
$\rho_{4,6\vert\cdot}$	-0.399	-0.387	-0.383	-0.339	-0.339	-0.296	-0.287	-0.267	-0.341	-0.339	1	1	0.45
$\rho_{5,7\vert\cdot}$	-0.202	-0.174	-0.166	-0.112	-0.113	-0.062	-0.057	-0.033	-0.111	-0.125	1	1	0.53
$\rho_{6,8\vert\cdot}$	-0.281	-0.260	-0.255	-0.209	-0.208	-0.162	-0.154	-0.112	-0.209	-0.216	1	1	0.54
$\rho_{9,11\vert\cdot}$	0.050	0.079	0.086	0.132	0.133	0.185	0.195	0.222	0.132	0.130	1	1	0.53
$\rho_{10,12\vert\cdot}$	-0.220	-0.198	-0.187	-0.141	-0.139	-0.089	-0.078	-0.064	-0.146	-0.131	1	1	0.53
$\rho_{1,4\vert\cdot}$	-0.283	-0.248	-0.238	-0.191	-0.191	-0.141	-0.131	-0.099	-0.194	-0.195	1	1	0.50
$\rho_{2,5\vert\cdot}$	0.045	0.079	0.085	0.144	0.142	0.192	0.201	0.216	0.145	0.136	1	1	0.54
$\rho_{4,7\vert\cdot}$	-0.332	-0.306	-0.295	-0.248	-0.249	-0.202	-0.192	-0.171	-0.249	-0.263	1	1	0.48
$\rho_{6,9\vert\cdot}$	-0.183	-0.150	-0.144	-0.090	-0.091	-0.037	-0.023	0.020	-0.089	-0.094	1	1	0.51
$\rho_{8,11\vert\cdot}$	0.298	0.313	0.326	0.379	0.377	0.420	0.427	0.442	0.381	0.378	1	1	0.49
$\rho_{1,5\vert\cdot}$	0.077	0.110	0.125	0.178	0.176	0.221	0.225	0.281	0.180	0.184	1	1	0.50
$\rho_{3,7\vert\cdot}$	-0.406	-0.381	-0.370	-0.330	-0.330	-0.288	-0.284	-0.270	-0.331	-0.329	1	1	0.50
$\rho_{8,12\vert\cdot}$	-0.208	-0.175	-0.168	-0.119	-0.118	-0.070	-0.057	-0.016	-0.117	-0.118	1	1	0.53
$\rho_{1,6\vert\cdot}$	-0.465	-0.434	-0.429	-0.389	-0.387	-0.343	-0.337	-0.308	-0.393	-0.400	1	1	0.48
$\rho_{2,7\vert\cdot}$	-0.182	-0.146	-0.138	-0.090	-0.090	-0.040	-0.032	-0.002	-0.090	-0.083	1	1	0.56
$\rho_{3,8\vert\cdot}$	-0.242	-0.227	-0.213	-0.161	-0.160	-0.112	-0.104	-0.039	-0.161	-0.170	1	1	0.54
$\rho_{4,9\vert\cdot}$	-0.466	-0.426	-0.423	-0.387	-0.385	-0.340	-0.332	-0.288	-0.390	-0.383	1	1	0.46
$\rho_{5,10\vert\cdot}$	0.035	0.079	0.085	0.139	0.138	0.185	0.193	0.212	0.139	0.135	1	1	0.53
$\rho_{6,11\vert\cdot}$	-0.237	-0.199	-0.191	-0.140	-0.139	-0.089	-0.081	-0.022	-0.141	-0.142	1	1	0.54
$\rho_{7,12\vert\cdot}$	-0.208	-0.174	-0.164	-0.116	-0.116	-0.063	-0.050	-0.041	-0.114	-0.113	1	1	0.53
$\rho_{1,7\vert\cdot}$	-0.176	-0.158	-0.142	-0.101	-0.101	-0.057	-0.050	-0.003	-0.100	-0.116	1	1	0.55
$\rho_{3,9\vert\cdot}$	-0.185	-0.169	-0.161	-0.114	-0.112	-0.063	-0.049	-0.021	-0.114	-0.126	1	1	0.53
$\rho_{4,10\vert\cdot}$	-0.206	-0.179	-0.173	-0.120	-0.120	-0.071	-0.060	-0.045	-0.120	-0.127	1	1	0.54
$\rho_{5,11\vert\cdot}$	-0.157	-0.121	-0.113	-0.061	-0.059	-0.005	0.009	0.034	-0.061	-0.061	1	0	0.56
$\rho_{6,12\vert\cdot}$	-0.247	-0.204	-0.193	-0.150	-0.148	-0.100	-0.095	-0.068	-0.151	-0.141	1	1	0.54
$\rho_{1,8\vert\cdot}$	-0.069	-0.017	-0.010	0.043	0.043	0.093	0.101	0.128	0.043	0.049	0	0	0.53
$\rho_{2,9\vert\cdot}$	0.001	0.047	0.062	0.111	0.110	0.157	0.165	0.188	0.111	0.103	1	1	0.54
$\rho_{3,10\vert\cdot}$	-0.292	-0.277	-0.269	-0.227	-0.225	-0.180	-0.172	-0.128	-0.229	-0.226	1	1	0.53
$\rho_{4,11\vert\cdot}$	-0.261	-0.219	-0.211	-0.162	-0.161	-0.110	-0.100	-0.066	-0.161	-0.164	1	1	0.56
$\rho_{5,12\vert\cdot}$	-0.273	-0.229	-0.220	-0.169	-0.168	-0.119	-0.110	-0.056	-0.169	-0.165	1	1	0.54
$\rho_{1,9\vert\cdot}$	0.255	0.290	0.298	0.345	0.345	0.390	0.401	0.417	0.345	0.347	1	1	0.50
$\rho_{2,10\vert\cdot}$	-0.272	-0.235	-0.221	-0.182	-0.180	-0.136	-0.130	-0.083	-0.183	-0.200	1	1	0.55
$\rho_{3,11\vert\cdot}$	-0.059	-0.018	-0.007	0.048	0.048	0.103	0.113	0.142	0.044	0.052	0	0	0.56
$\rho_{4,12\vert\cdot}$	-0.353	-0.323	-0.305	-0.257	-0.259	-0.213	-0.201	-0.163	-0.256	-0.255	1	1	0.53
$\rho_{1,10\vert\cdot}$	0.169	0.209	0.221	0.267	0.267	0.312	0.321	0.339	0.269	0.262	1	1	0.54
$\rho_{3,12\vert\cdot}$	0.186	0.203	0.210	0.258	0.259	0.308	0.319	0.344	0.258	0.262	1	1	0.54
$\rho_{2,12\vert\cdot}$	-0.255	-0.207	-0.202	-0.150	-0.150	-0.103	-0.097	-0.058	-0.149	-0.149	1	1	0.56

Table B.13: MCMC results for the copula parameters of the reduced model of the 1st vine construction with credible level 10%

θ	$\hat{\theta}_{min}$	$\hat{\theta}_{2.5\%}$	$\hat{\theta}_{5\%}$	$\hat{\theta}_{med}$	$\bar{\theta}$	$\hat{\theta}_{95\%}$	$\hat{\theta}_{97.5\%}$	$\hat{\theta}_{max}$	$\hat{\theta}_{mod}$	$\hat{\theta}_{IFM}$	$C_{10\%}$	$C_{5\%}$	$\bar{p}_{acc}$
$\rho_{2,3}$	0.135	0.195	0.205	0.255	0.254	0.305	0.310	0.333	0.257	0.260	1	1	0.52
$\rho_{3,4}$	-0.136	-0.124	-0.114	-0.065	-0.064	-0.014	-0.003	0.019	-0.066	-0.066	1	1	0.51
$\rho_{4,5}$	-0.151	-0.112	-0.107	-0.061	-0.059	-0.010	0.003	0.025	-0.063	-0.066	1	0	0.48
$\rho_{5,6}$	-0.193	-0.175	-0.168	-0.120	-0.121	-0.074	-0.060	0.002	-0.120	-0.123	1	1	0.52
$\rho_{6,7}$	-0.141	-0.114	-0.108	-0.058	-0.057	-0.009	0.001	0.035	-0.058	-0.060	1	0	0.52
$\rho_{7,8}$	0.013	0.073	0.082	0.131	0.130	0.177	0.186	0.235	0.134	0.132	1	1	0.55
$\rho_{9,10}$	0.012	0.054	0.067	0.114	0.113	0.162	0.172	0.205	0.112	0.114	1	1	0.54
$\rho_{10,11}$	-0.010	0.015	0.028	0.073	0.073	0.115	0.126	0.159	0.075	0.090	1	1	0.53
$\rho_{11,12}$	-0.005	0.036	0.048	0.095	0.094	0.142	0.150	0.164	0.095	0.103	1	1	0.53
$\rho_{2,4\vert\cdot}$	-0.249	-0.226	-0.218	-0.165	-0.165	-0.113	-0.104	-0.087	-0.167	-0.162	1	1	0.53
$\rho_{3,5\vert\cdot}$	0.134	0.186	0.193	0.241	0.240	0.286	0.294	0.320	0.240	0.243	1	1	0.50
$\rho_{4,6\vert\cdot}$	-0.399	-0.380	-0.374	-0.326	-0.329	-0.281	-0.276	-0.250	-0.324	-0.339	1	1	0.45
$\rho_{5,7\vert\cdot}$	-0.225	-0.180	-0.167	-0.117	-0.117	-0.066	-0.059	-0.026	-0.115	-0.125	1	1	0.52
$\rho_{6,8\vert\cdot}$	-0.310	-0.272	-0.266	-0.215	-0.215	-0.165	-0.149	-0.106	-0.214	-0.216	1	1	0.54
$\rho_{9,11\vert\cdot}$	0.062	0.096	0.104	0.152	0.152	0.200	0.207	0.232	0.150	0.130	1	1	0.53
$\rho_{10,12\vert\cdot}$	-0.219	-0.199	-0.191	-0.134	-0.134	-0.082	-0.073	-0.042	-0.133	-0.131	1	1	0.52
$\rho_{1,4\vert\cdot}$	-0.274	-0.245	-0.241	-0.194	-0.192	-0.142	-0.134	-0.105	-0.195	-0.195	1	1	0.48
$\rho_{2,5\vert\cdot}$	0.049	0.075	0.084	0.141	0.140	0.190	0.196	0.229	0.143	0.136	1	1	0.55
$\rho_{4,7\vert\cdot}$	-0.330	-0.304	-0.293	-0.249	-0.248	-0.204	-0.198	-0.184	-0.251	-0.263	1	1	0.49
$\rho_{6,9\vert\cdot}$	-0.216	-0.159	-0.148	-0.098	-0.098	-0.051	-0.038	-0.012	-0.098	-0.094	1	1	0.51
$\rho_{8,11\vert\cdot}$	0.292	0.323	0.329	0.375	0.373	0.414	0.419	0.439	0.377	0.378	1	1	0.51
$\rho_{1,5\vert\cdot}$	0.113	0.127	0.133	0.177	0.178	0.229	0.238	0.272	0.176	0.184	1	1	0.50
$\rho_{3,7\vert\cdot}$	-0.397	-0.380	-0.370	-0.329	-0.328	-0.283	-0.277	-0.268	-0.332	-0.329	1	1	0.50
$\rho_{8,12\vert\cdot}$	-0.217	-0.183	-0.167	-0.118	-0.118	-0.068	-0.061	-0.044	-0.119	-0.118	1	1	0.53
$\rho_{1,6\vert\cdot}$	-0.475	-0.442	-0.437	-0.393	-0.393	-0.349	-0.334	-0.303	-0.394	-0.400	1	1	0.47
$\rho_{2,7\vert\cdot}$	-0.174	-0.155	-0.143	-0.090	-0.090	-0.035	-0.023	-0.009	-0.089	-0.083	1	1	0.56
$\rho_{3,8\vert\cdot}$	-0.254	-0.234	-0.226	-0.173	-0.173	-0.123	-0.112	-0.056	-0.173	-0.170	1	1	0.54
$\rho_{4,9\vert\cdot}$	-0.457	-0.437	-0.428	-0.383	-0.384	-0.339	-0.331	-0.278	-0.383	-0.383	1	1	0.46
$\rho_{5,10\vert\cdot}$	0.060	0.076	0.086	0.132	0.132	0.181	0.190	0.219	0.130	0.135	1	1	0.54
$\rho_{6,11\vert\cdot}$	-0.233	-0.197	-0.191	-0.144	-0.141	-0.086	-0.077	-0.059	-0.149	-0.142	1	1	0.57
$\rho_{7,12\vert\cdot}$	-0.196	-0.171	-0.158	-0.109	-0.110	-0.063	-0.056	-0.029	-0.108	-0.113	1	1	0.54
$\rho_{1,7\vert\cdot}$	-0.194	-0.161	-0.153	-0.106	-0.105	-0.054	-0.047	-0.024	-0.106	-0.116	1	1	0.54
$\rho_{3,9\vert\cdot}$	-0.213	-0.172	-0.163	-0.117	-0.116	-0.066	-0.060	-0.044	-0.119	-0.126	1	1	0.53
$\rho_{4,10\vert\cdot}$	-0.210	-0.177	-0.169	-0.115	-0.117	-0.066	-0.063	-0.007	-0.115	-0.127	1	1	0.52
$\rho_{5,11\vert\cdot}$	-0.143	-0.119	-0.114	-0.062	-0.061	-0.009	-0.001	0.031	-0.063	-0.061	1	1	0.57
$\rho_{6,12\vert\cdot}$	-0.231	-0.209	-0.204	-0.155	-0.154	-0.106	-0.098	-0.046	-0.157	-0.141	1	1	0.53
$\rho_{2,9\vert\cdot}$	0.013	0.050	0.061	0.107	0.108	0.164	0.171	0.197	0.104	0.103	1	1	0.54
$\rho_{3,10\vert\cdot}$	-0.297	-0.276	-0.269	-0.224	-0.225	-0.179	-0.172	-0.147	-0.226	-0.226	1	1	0.53
$\rho_{4,11\vert\cdot}$	-0.251	-0.225	-0.216	-0.157	-0.159	-0.104	-0.097	-0.046	-0.156	-0.164	1	1	0.55
$\rho_{5,12\vert\cdot}$	-0.241	-0.230	-0.217	-0.169	-0.170	-0.120	-0.113	-0.066	-0.166	-0.165	1	1	0.54
$\rho_{1,9\vert\cdot}$	0.273	0.291	0.298	0.347	0.345	0.386	0.395	0.408	0.350	0.347	1	1	0.50
$\rho_{2,10\vert\cdot}$	-0.264	-0.238	-0.227	-0.181	-0.180	-0.131	-0.124	-0.091	-0.183	-0.200	1	1	0.54
$\rho_{4,12\vert\cdot}$	-0.349	-0.318	-0.312	-0.259	-0.259	-0.211	-0.206	-0.186	-0.259	-0.255	1	1	0.53
$\rho_{1,10\vert\cdot}$	0.176	0.203	0.218	0.270	0.268	0.318	0.325	0.348	0.270	0.262	1	1	0.53
$\rho_{3,12\vert\cdot}$	0.182	0.203	0.212	0.257	0.258	0.306	0.313	0.351	0.255	0.262	1	1	0.54
$\rho_{2,12\vert\cdot}$	-0.273	-0.207	-0.197	-0.150	-0.150	-0.103	-0.092	-0.062	-0.151	-0.149	1	1	0.57

Table B.14: MCMC results for the copula parameters of the reduced model of the 1st vine construction with credible level 5%

θ	$\hat{\theta}_{min}$	$\hat{\theta}_{2.5\%}$	$\hat{\theta}_{5\%}$	$\hat{\theta}_{med}$	$\bar{\theta}$	$\hat{\theta}_{95\%}$	$\hat{\theta}_{97.5\%}$	$\hat{\theta}_{max}$	$\hat{\theta}_{mod}$	$\hat{\theta}_{IFM}$	$\bar{p}_{acc}$
$\beta_1(=\beta_U)$	0.756	0.769	0.773	0.806	0.807	0.841	0.844	0.869	0.806	0.805	1.000
$\beta_2(=\beta_\$)$	1.115	1.129	1.134	1.159	1.159	1.183	1.187	1.205	1.160	1.158	1.000
$\beta_3(=\beta_B)$	1.246	1.255	1.262	1.295	1.293	1.321	1.327	1.348	1.297	1.294	1.000
$\beta_4(=\beta_E)$	0.793	0.812	0.818	0.856	0.857	0.895	0.900	0.925	0.856	0.858	1.000
$\beta_5(=\beta_S)$	0.914	0.927	0.933	0.960	0.961	0.990	0.995	1.013	0.960	0.961	1.000
$\beta_6(=\beta_N)$	0.730	0.741	0.742	0.768	0.767	0.787	0.793	0.806	0.769	0.767	1.000
$\beta_7(=\beta_H)$	0.808	0.825	0.833	0.864	0.864	0.902	0.907	0.915	0.862	0.865	1.000
$\beta_8(=\beta_D)$	1.156	1.179	1.183	1.218	1.218	1.255	1.263	1.279	1.217	1.218	1.000
$\beta_9(=\beta_C)$	0.941	0.953	0.956	0.978	0.979	1.002	1.007	1.016	0.977	0.978	1.000
$\beta_{10}(=\beta_O)$	1.093	1.107	1.112	1.135	1.136	1.161	1.167	1.173	1.134	1.135	1.000
$\beta_{11}(=\beta_M)$	1.162	1.167	1.171	1.189	1.189	1.208	1.211	1.227	1.189	1.189	1.000
$\beta_{12}(=\beta_T)$	0.589	0.606	0.615	0.645	0.644	0.673	0.676	0.711	0.644	0.643	1.000
$\sigma_1^2(=\sigma_U^2)$	11.840	12.380	12.530	13.476	13.506	14.624	14.723	15.040	13.425	13.456	0.298
$\sigma_2^2(=\sigma_\$^2)$	6.640	6.927	6.967	7.536	7.559	8.180	8.262	8.575	7.526	7.527	0.306
$\sigma_3^2(=\sigma_B^2)$	9.381	9.631	9.794	10.560	10.571	11.383	11.527	12.359	10.506	10.521	0.284
$\sigma_4^2(=\sigma_E^2)$	12.618	13.243	13.403	14.564	14.572	15.794	16.003	16.468	14.549	14.412	0.282
$\sigma_5^2(=\sigma_S^2)$	6.829	7.177	7.278	7.819	7.843	8.433	8.613	8.967	7.808	7.829	0.294
$\sigma_6^2(=\sigma_N^2)$	4.346	4.612	4.657	4.978	5.017	5.439	5.514	5.902	4.960	5.017	0.277
$\sigma_7^2(=\sigma_H^2)$	10.207	10.742	10.930	11.802	11.824	12.791	12.986	13.666	11.787	11.803	0.295
$\sigma_8^2(=\sigma_D^2)$	12.610	13.121	13.411	14.447	14.441	15.567	15.760	16.637	14.465	14.427	0.301
$\sigma_9^2(=\sigma_C^2)$	5.881	6.047	6.111	6.554	6.571	7.019	7.100	7.661	6.555	6.548	0.295
$\sigma_{10}^2(=\sigma_O^2)$	6.048	6.296	6.355	6.836	6.827	7.325	7.380	7.896	6.833	6.790	0.295
$\sigma_{11}^2(=\sigma_M^2)$	3.568	3.708	3.772	4.084	4.094	4.449	4.511	4.796	4.081	4.083	0.285
$\sigma_{12}^2(=\sigma_T^2)$	7.983	8.223	8.333	8.926	8.953	9.634	9.810	10.143	8.882	8.942	0.304

Table B.15: MCMC results for the marginal parameters of the reduced model of the 2nd vine construction with credible level 10%.

θ	$\hat{\theta}_{min}$	$\hat{\theta}_{2.5\%}$	$\hat{\theta}_{5\%}$	$\hat{\theta}_{med}$	$\bar{\theta}$	$\hat{\theta}_{95\%}$	$\hat{\theta}_{97.5\%}$	$\hat{\theta}_{max}$	$\hat{\theta}_{mod}$	$\hat{\theta}_{IFM}$	$\bar{p}_{acc}$
$\beta_1(=\beta_U)$	0.742	0.764	0.769	0.805	0.804	0.841	0.846	0.863	0.806	0.805	1.000
$\beta_2(=\beta_\$)$	1.108	1.126	1.132	1.158	1.158	1.186	1.192	1.204	1.159	1.158	1.000
$\beta_3(=\beta_B)$	1.240	1.259	1.265	1.294	1.294	1.327	1.332	1.358	1.295	1.294	1.000
$\beta_4(=\beta_E)$	0.793	0.816	0.823	0.859	0.859	0.896	0.903	0.926	0.859	0.858	1.000
$\beta_5(=\beta_S)$	0.920	0.930	0.934	0.961	0.961	0.986	0.995	1.014	0.962	0.961	1.000
$\beta_6(=\beta_N)$	0.730	0.741	0.748	0.767	0.768	0.789	0.794	0.807	0.767	0.767	1.000
$\beta_7(=\beta_H)$	0.808	0.831	0.838	0.867	0.867	0.899	0.905	0.927	0.868	0.865	1.000
$\beta_8(=\beta_D)$	1.144	1.174	1.182	1.218	1.217	1.253	1.261	1.274	1.219	1.218	1.000
$\beta_9(=\beta_C)$	0.931	0.948	0.954	0.980	0.979	1.006	1.012	1.032	0.981	0.978	1.000
$\beta_{10}(=\beta_O)$	1.093	1.105	1.110	1.136	1.135	1.159	1.165	1.184	1.135	1.135	1.000
$\beta_{11}(=\beta_M)$	1.160	1.167	1.171	1.191	1.190	1.208	1.209	1.227	1.193	1.189	1.000
$\beta_{12}(=\beta_T)$	0.573	0.606	0.612	0.643	0.643	0.676	0.683	0.697	0.643	0.643	1.000
$\sigma_1^2(=\sigma_U^2)$	11.921	12.353	12.517	13.570	13.558	14.597	14.772	15.941	13.634	13.456	0.295
$\sigma_2^2(=\sigma_\$^2)$	6.668	6.917	7.039	7.485	7.517	8.099	8.216	8.616	7.477	7.527	0.296
$\sigma_3^2(=\sigma_B^2)$	9.337	9.739	9.943	10.691	10.684	11.448	11.581	12.031	10.688	10.521	0.281
$\sigma_4^2(=\sigma_E^2)$	12.833	13.359	13.539	14.472	14.510	15.619	15.855	16.436	14.443	14.412	0.271
$\sigma_5^2(=\sigma_S^2)$	6.948	7.073	7.165	7.781	7.784	8.330	8.458	8.916	7.792	7.829	0.295
$\sigma_6^2(=\sigma_N^2)$	4.328	4.578	4.616	5.012	5.000	5.389	5.453	5.551	5.018	5.017	0.288
$\sigma_7^2(=\sigma_H^2)$	10.176	10.693	10.896	11.791	11.804	12.759	12.954	13.606	11.794	11.803	0.307
$\sigma_8^2(=\sigma_D^2)$	12.622	13.375	13.566	14.432	14.483	15.531	15.764	16.566	14.363	14.427	0.305
$\sigma_9^2(=\sigma_C^2)$	5.666	5.982	6.086	6.568	6.570	7.122	7.235	7.524	6.546	6.548	0.296
$\sigma_{10}^2(=\sigma_O^2)$	5.955	6.250	6.320	6.792	6.820	7.309	7.357	7.709	6.797	6.790	0.293
$\sigma_{11}^2(=\sigma_M^2)$	3.615	3.788	3.808	4.090	4.093	4.417	4.476	4.634	4.080	4.083	0.295
$\sigma_{12}^2(=\sigma_T^2)$	7.770	8.174	8.264	9.011	8.993	9.642	9.775	10.204	9.021	8.942	0.295

Table B.16: MCMC results for the marginal parameters of the reduced model of the 2nd vine construction with credible level 5%.

θ	$\hat{\theta}_{min}$	$\hat{\theta}_{2.5\%}$	$\hat{\theta}_{5\%}$	$\hat{\theta}_{med}$	$\bar{\theta}$	$\hat{\theta}_{95\%}$	$\hat{\theta}_{97.5\%}$	$\hat{\theta}_{max}$	$\hat{\theta}_{mod}$	$\hat{\theta}_{IFM}$	$C_{10\%}$	$C_{5\%}$	$\bar{p}_{acc}$
$\rho_{1,2}$	0.273	0.300	0.312	0.359	0.359	0.406	0.411	0.449	0.360	0.367	1	1	0.49
$\rho_{2,3}$	-0.314	-0.277	-0.267	-0.219	-0.220	-0.174	-0.164	-0.146	-0.218	-0.224	1	1	0.54
$\rho_{3,4}$	-0.413	-0.380	-0.370	-0.326	-0.327	-0.285	-0.279	-0.252	-0.327	-0.328	1	1	0.48
$\rho_{4,5}$	-0.396	-0.373	-0.364	-0.315	-0.316	-0.271	-0.265	-0.236	-0.314	-0.324	1	1	0.49
$\rho_{5,6}$	0.329	0.354	0.367	0.406	0.406	0.448	0.454	0.500	0.407	0.414	1	1	0.46
$\rho_{6,7}$	0.233	0.274	0.286	0.333	0.331	0.377	0.383	0.410	0.334	0.332	1	1	0.52
$\rho_{7,8}$	-0.225	-0.213	-0.206	-0.152	-0.152	-0.100	-0.086	-0.058	-0.152	-0.156	1	1	0.55
$\rho_{8,9}$	0.080	0.132	0.138	0.188	0.189	0.243	0.254	0.290	0.186	0.197	1	1	0.53
$\rho_{9,10}$	-0.221	-0.208	-0.200	-0.153	-0.151	-0.101	-0.087	-0.061	-0.157	-0.147	1	1	0.54
$\rho_{10,11}$	0.205	0.244	0.255	0.295	0.296	0.341	0.351	0.392	0.295	0.294	1	1	0.50
$\rho_{11,12}$	-0.412	-0.382	-0.374	-0.325	-0.325	-0.277	-0.269	-0.249	-0.326	-0.325	1	1	0.50
$\rho_{1,3\vert\cdot}$	-0.227	-0.206	-0.198	-0.151	-0.151	-0.096	-0.087	-0.066	-0.153	-0.154	1	1	0.55
$\rho_{2,4\vert\cdot}$	-0.272	-0.251	-0.243	-0.189	-0.191	-0.140	-0.128	-0.101	-0.186	-0.193	1	1	0.56
$\rho_{3,5\vert\cdot}$	-0.275	-0.252	-0.247	-0.198	-0.199	-0.152	-0.144	-0.099	-0.196	-0.208	1	1	0.51
$\rho_{4,6\vert\cdot}$	-0.131	-0.103	-0.089	-0.041	-0.040	0.010	0.017	0.061	-0.042	-0.061	0	0	0.53
$\rho_{7,9\vert\cdot}$	0.093	0.124	0.132	0.178	0.179	0.231	0.238	0.280	0.178	0.178	1	1	0.54
$\rho_{9,11\vert\cdot}$	0.195	0.246	0.253	0.302	0.301	0.342	0.346	0.387	0.301	0.306	1	1	0.53
$\rho_{1,4\vert\cdot}$	-0.010	0.001	0.010	0.058	0.058	0.107	0.114	0.131	0.057	0.059	1	1	0.56
$\rho_{2,5\vert\cdot}$	-0.028	0.005	0.009	0.058	0.058	0.112	0.126	0.152	0.056	0.058	1	1	0.56
$\rho_{3,6\vert\cdot}$	-0.455	-0.402	-0.394	-0.348	-0.349	-0.305	-0.295	-0.270	-0.347	-0.367	1	1	0.51
$\rho_{4,7\vert\cdot}$	-0.206	-0.158	-0.151	-0.096	-0.095	-0.046	-0.041	0.010	-0.092	-0.097	1	1	0.54
$\rho_{5,8\vert\cdot}$	0.085	0.113	0.123	0.168	0.170	0.217	0.226	0.257	0.169	0.175	1	1	0.55
$\rho_{6,9\vert\cdot}$	0.035	0.076	0.085	0.132	0.133	0.183	0.191	0.218	0.131	0.133	1	1	0.57
$\rho_{7,10\vert\cdot}$	-0.194	-0.155	-0.147	-0.106	-0.105	-0.058	-0.049	-0.015	-0.107	-0.108	1	1	0.54
$\rho_{8,11\vert\cdot}$	0.169	0.186	0.193	0.239	0.240	0.289	0.297	0.324	0.237	0.244	1	1	0.54
$\rho_{1,5\vert\cdot}$	-0.215	-0.174	-0.164	-0.112	-0.111	-0.059	-0.052	-0.018	-0.113	-0.111	1	1	0.56
$\rho_{2,6\vert\cdot}$	-0.002	0.021	0.034	0.083	0.083	0.135	0.151	0.171	0.082	0.082	1	1	0.58
$\rho_{4,8\vert\cdot}$	-0.247	-0.228	-0.218	-0.165	-0.166	-0.115	-0.105	-0.062	-0.164	-0.159	1	1	0.54
$\rho_{5,9\vert\cdot}$	-0.191	-0.148	-0.141	-0.088	-0.089	-0.035	-0.028	-0.002	-0.086	-0.089	1	1	0.57
$\rho_{6,10\vert\cdot}$	-0.036	-0.006	-0.001	0.056	0.056	0.110	0.118	0.135	0.057	0.067	0	0	0.56
$\rho_{7,11\vert\cdot}$	-0.222	-0.186	-0.181	-0.131	-0.129	-0.079	-0.067	-0.039	-0.129	-0.132	1	1	0.56
$\rho_{1,6\vert\cdot}$	-0.022	0.001	0.013	0.062	0.061	0.109	0.119	0.164	0.062	0.060	1	1	0.56
$\rho_{3,8\vert\cdot}$	-0.160	-0.143	-0.133	-0.086	-0.087	-0.038	-0.032	-0.009	-0.086	-0.078	1	1	0.55
$\rho_{5,10\vert\cdot}$	-0.142	-0.123	-0.112	-0.059	-0.058	-0.006	0.004	0.023	-0.060	-0.061	1	0	0.56
$\rho_{7,12\vert\cdot}$	-0.143	-0.129	-0.122	-0.067	-0.067	-0.019	-0.008	0.016	-0.065	-0.078	1	1	0.55
$\rho_{3,9\vert\cdot}$	-0.201	-0.145	-0.143	-0.093	-0.094	-0.045	-0.035	-0.003	-0.094	-0.088	1	1	0.56
$\rho_{4,10\vert\cdot}$	-0.213	-0.199	-0.196	-0.152	-0.151	-0.102	-0.090	-0.066	-0.153	-0.138	1	1	0.54
$\rho_{5,11\vert\cdot}$	-0.131	-0.125	-0.113	-0.063	-0.063	-0.012	-0.006	0.027	-0.062	-0.063	1	1	0.56
$\rho_{1,8\vert\cdot}$	-0.143	-0.114	-0.102	-0.056	-0.055	-0.005	0.000	0.024	-0.058	-0.059	1	0	0.57
$\rho_{2,9\vert\cdot}$	-0.184	-0.150	-0.143	-0.092	-0.092	-0.045	-0.035	-0.010	-0.092	-0.093	1	1	0.57
$\rho_{3,10\vert\cdot}$	-0.297	-0.276	-0.269	-0.220	-0.221	-0.170	-0.161	-0.142	-0.219	-0.214	1	1	0.53
$\rho_{2,10\vert\cdot}$	-0.033	-0.011	-0.001	0.053	0.052	0.101	0.107	0.142	0.053	0.053	0	0	0.57
$\rho_{3,11\vert\cdot}$	0.008	0.032	0.046	0.100	0.098	0.149	0.154	0.195	0.102	0.095	1	1	0.56
$\rho_{4,12\vert\cdot}$	-0.366	-0.341	-0.331	-0.279	-0.280	-0.230	-0.224	-0.210	-0.277	-0.292	1	1	0.53
$\rho_{1,10\vert\cdot}$	-0.234	-0.213	-0.205	-0.147	-0.147	-0.094	-0.081	-0.042	-0.146	-0.148	1	1	0.56
$\rho_{3,12\vert\cdot}$	-0.309	-0.281	-0.270	-0.222	-0.221	-0.174	-0.168	-0.138	-0.220	-0.228	1	1	0.55
$\rho_{1,11\vert\cdot}$	-0.241	-0.221	-0.215	-0.170	-0.169	-0.116	-0.110	-0.092	-0.173	-0.170	1	1	0.56

Table B.17: MCMC results for the copula parameters of the reduced model of the 2nd vine construction with credible level 10%.

θ	$\hat{\theta}_{min}$	$\hat{\theta}_{2.5\%}$	$\hat{\theta}_{5\%}$	$\hat{\theta}_{med}$	$\bar{\theta}$	$\hat{\theta}_{95\%}$	$\hat{\theta}_{97.5\%}$	$\hat{\theta}_{max}$	$\hat{\theta}_{mod}$	$\hat{\theta}_{IFM}$	$C_{10\%}$	$C_{5\%}$	$\bar{p}_{acc}$
$\rho_{1,2}$	0.283	0.315	0.320	0.366	0.367	0.413	0.419	0.458	0.365	0.367	1	1	0.50
$\rho_{2,3}$	-0.295	-0.277	-0.266	-0.218	-0.217	-0.167	-0.157	-0.116	-0.216	-0.224	1	1	0.55
$\rho_{3,4}$	-0.405	-0.384	-0.380	-0.340	-0.339	-0.296	-0.284	-0.255	-0.339	-0.328	1	1	0.48
$\rho_{4,5}$	-0.393	-0.362	-0.357	-0.313	-0.312	-0.261	-0.257	-0.226	-0.315	-0.324	1	1	0.50
$\rho_{5,6}$	0.330	0.352	0.364	0.403	0.403	0.445	0.453	0.468	0.403	0.414	1	1	0.46
$\rho_{6,7}$	0.236	0.263	0.275	0.331	0.328	0.375	0.382	0.418	0.333	0.332	1	1	0.52
$\rho_{7,8}$	-0.239	-0.219	-0.205	-0.156	-0.156	-0.108	-0.097	-0.043	-0.155	-0.156	1	1	0.54
$\rho_{8,9}$	0.095	0.136	0.143	0.189	0.190	0.238	0.247	0.283	0.189	0.197	1	1	0.54
$\rho_{9,10}$	-0.250	-0.213	-0.202	-0.158	-0.156	-0.109	-0.096	-0.063	-0.160	-0.147	1	1	0.52
$\rho_{10,11}$	0.218	0.244	0.252	0.294	0.294	0.340	0.348	0.374	0.296	0.294	1	1	0.50
$\rho_{11,12}$	-0.405	-0.374	-0.367	-0.323	-0.324	-0.284	-0.273	-0.208	-0.324	-0.325	1	1	0.50
$\rho_{1,3\vert\cdot}$	-0.256	-0.206	-0.196	-0.148	-0.150	-0.103	-0.099	-0.071	-0.148	-0.154	1	1	0.55
$\rho_{2,4\vert\cdot}$	-0.271	-0.240	-0.231	-0.186	-0.186	-0.134	-0.125	-0.093	-0.187	-0.193	1	1	0.55
$\rho_{3,5\vert\cdot}$	-0.290	-0.260	-0.249	-0.200	-0.201	-0.156	-0.144	-0.133	-0.199	-0.208	1	1	0.52
$\rho_{7,9\vert\cdot}$	0.078	0.127	0.134	0.180	0.180	0.232	0.243	0.269	0.180	0.178	1	1	0.55
$\rho_{9,11\vert\cdot}$	0.212	0.251	0.255	0.297	0.299	0.347	0.357	0.382	0.294	0.306	1	1	0.52
$\rho_{3,6\vert\cdot}$	-0.437	-0.410	-0.396	-0.354	-0.353	-0.307	-0.295	-0.268	-0.357	-0.367	1	1	0.49
$\rho_{4,7\vert\cdot}$	-0.185	-0.156	-0.149	-0.095	-0.095	-0.040	-0.028	0.018	-0.094	-0.097	1	1	0.55
$\rho_{5,8\vert\cdot}$	0.075	0.115	0.122	0.173	0.173	0.225	0.235	0.266	0.171	0.175	1	1	0.56
$\rho_{6,9\vert\cdot}$	0.038	0.073	0.081	0.133	0.133	0.183	0.190	0.202	0.132	0.133	1	1	0.57
$\rho_{7,10\vert\cdot}$	-0.190	-0.166	-0.159	-0.111	-0.111	-0.063	-0.052	-0.032	-0.111	-0.108	1	1	0.56
$\rho_{8,11\vert\cdot}$	0.153	0.179	0.190	0.240	0.240	0.289	0.297	0.320	0.241	0.244	1	1	0.54
$\rho_{1,5\vert\cdot}$	-0.231	-0.166	-0.160	-0.106	-0.108	-0.059	-0.049	-0.026	-0.101	-0.111	1	1	0.57
$\rho_{2,6\vert\cdot}$	-0.004	0.014	0.024	0.079	0.079	0.135	0.146	0.166	0.078	0.082	1	1	0.57
$\rho_{4,8\vert\cdot}$	-0.260	-0.225	-0.215	-0.164	-0.165	-0.115	-0.103	-0.053	-0.165	-0.159	1	1	0.55
$\rho_{5,9\vert\cdot}$	-0.158	-0.143	-0.135	-0.079	-0.081	-0.031	-0.015	0.011	-0.077	-0.089	1	1	0.57
$\rho_{6,10\vert\cdot}$	-0.037	-0.002	0.010	0.054	0.056	0.106	0.117	0.139	0.049	0.067	1	0	0.56
$\rho_{7,11\vert\cdot}$	-0.209	-0.189	-0.182	-0.132	-0.133	-0.080	-0.076	-0.043	-0.129	-0.132	1	1	0.55
$\rho_{1,6\vert\cdot}$	-0.018	0.004	0.017	0.059	0.062	0.114	0.125	0.155	0.054	0.060	1	1	0.57
$\rho_{3,8\vert\cdot}$	-0.164	-0.143	-0.133	-0.083	-0.085	-0.039	-0.033	0.012	-0.082	-0.078	1	1	0.56
$\rho_{5,10\vert\cdot}$	-0.140	-0.129	-0.117	-0.061	-0.061	-0.010	-0.003	0.030	-0.058	-0.061	1	1	0.55
$\rho_{7,12\vert\cdot}$	-0.148	-0.130	-0.121	-0.072	-0.071	-0.021	-0.009	0.023	-0.071	-0.078	1	1	0.56
$\rho_{3,9\vert\cdot}$	-0.169	-0.146	-0.137	-0.089	-0.090	-0.041	-0.032	-0.009	-0.090	-0.088	1	1	0.55
$\rho_{4,10\vert\cdot}$	-0.265	-0.202	-0.195	-0.144	-0.144	-0.096	-0.088	-0.066	-0.144	-0.138	1	1	0.54
$\rho_{5,11\vert\cdot}$	-0.155	-0.128	-0.118	-0.065	-0.067	-0.018	-0.010	0.021	-0.061	-0.063	1	1	0.58
$\rho_{1,8\vert\cdot}$	-0.136	-0.117	-0.104	-0.056	-0.055	-0.005	0.007	0.034	-0.056	-0.059	1	0	0.57
$\rho_{2,9\vert\cdot}$	-0.191	-0.149	-0.142	-0.091	-0.091	-0.041	-0.028	0.004	-0.092	-0.093	1	1	0.58
$\rho_{3,10\vert\cdot}$	-0.320	-0.271	-0.264	-0.216	-0.217	-0.167	-0.155	-0.129	-0.217	-0.214	1	1	0.53
$\rho_{3,11\vert\cdot}$	0.013	0.042	0.054	0.102	0.102	0.153	0.164	0.187	0.100	0.095	1	1	0.56
$\rho_{4,12\vert\cdot}$	-0.373	-0.346	-0.338	-0.287	-0.288	-0.236	-0.226	-0.187	-0.286	-0.292	1	1	0.53
$\rho_{1,10\vert\cdot}$	-0.229	-0.208	-0.197	-0.145	-0.143	-0.086	-0.076	-0.039	-0.146	-0.148	1	1	0.56
$\rho_{3,12\vert\cdot}$	-0.292	-0.278	-0.272	-0.223	-0.222	-0.168	-0.160	-0.136	-0.224	-0.228	1	1	0.55
$\rho_{1,11\vert\cdot}$	-0.261	-0.224	-0.215	-0.169	-0.168	-0.118	-0.107	-0.068	-0.171	-0.170	1	1	0.56

Table B.18: MCMC results for the copula parameters of the reduced model of the 2nd vine construction with credible level 5%

θ	$\hat{\theta}_{min}$	$\hat{\theta}_{2.5\%}$	$\hat{\theta}_{5\%}$	$\hat{\theta}_{med}$	$\bar{\theta}$	$\hat{\theta}_{95\%}$	$\hat{\theta}_{97.5\%}$	$\hat{\theta}_{max}$	$\hat{\theta}_{mod}$	$\hat{\theta}_{IFM}$	$\bar{p}_{acc}$
$\beta_1(=\beta_B)$	1.241	1.260	1.263	1.295	1.294	1.322	1.328	1.349	1.295	1.294	1.000
$\beta_2(=\beta_\$)$	1.102	1.129	1.131	1.158	1.158	1.183	1.187	1.195	1.157	1.158	1.000
$\beta_3(=\beta_N)$	0.730	0.742	0.747	0.765	0.766	0.787	0.793	0.807	0.765	0.767	1.000
$\beta_4(=\beta_E)$	0.776	0.816	0.824	0.858	0.857	0.891	0.897	0.913	0.856	0.858	1.000
$\beta_5(=\beta_D)$	1.159	1.178	1.185	1.220	1.220	1.254	1.260	1.274	1.221	1.218	1.000
$\beta_6(=\beta_T)$	0.587	0.610	0.618	0.643	0.643	0.672	0.678	0.699	0.641	0.643	1.000
$\beta_7(=\beta_O)$	1.091	1.103	1.109	1.134	1.135	1.159	1.162	1.177	1.134	1.135	1.000
$\beta_8(=\beta_C)$	0.929	0.947	0.952	0.978	0.978	1.001	1.005	1.025	0.977	0.978	1.000
$\beta_9(=\beta_U)$	0.749	0.766	0.770	0.803	0.805	0.840	0.845	0.864	0.803	0.805	1.000
$\beta_{10}(=\beta_S)$	0.914	0.929	0.937	0.963	0.963	0.990	0.995	1.015	0.963	0.961	1.000
$\beta_{11}(=\beta_H)$	0.818	0.828	0.833	0.863	0.864	0.898	0.907	0.922	0.863	0.865	1.000
$\beta_{12}(=\beta_M)$	1.151	1.168	1.170	1.190	1.190	1.210	1.215	1.219	1.190	1.189	1.000
$\sigma^2_1(=\sigma^2_B)$	9.078	9.739	9.921	10.575	10.619	11.430	11.544	11.948	10.500	10.521	0.290
$\sigma^2_2(=\sigma^2_\$)$	6.658	6.920	6.980	7.547	7.522	8.046	8.124	8.587	7.584	7.527	0.291
$\sigma^2_3(=\sigma^2_N)$	4.370	4.573	4.646	4.965	4.961	5.323	5.367	5.656	4.954	5.017	0.278
$\sigma^2_4(=\sigma^2_E)$	12.946	13.363	13.498	14.439	14.463	15.554	15.662	16.582	14.413	14.412	0.276
$\sigma^2_5(=\sigma^2_D)$	12.807	13.170	13.381	14.428	14.491	15.626	15.929	16.502	14.342	14.427	0.299
$\sigma^2_6(=\sigma^2_T)$	8.231	8.346	8.446	9.001	9.025	9.644	9.791	10.359	8.981	8.942	0.294
$\sigma^2_7(=\sigma^2_O)$	5.947	6.275	6.333	6.813	6.820	7.319	7.499	7.810	6.796	6.790	0.302
$\sigma^2_8(=\sigma^2_C)$	5.801	6.005	6.092	6.568	6.570	7.059	7.134	7.540	6.555	6.548	0.300
$\sigma^2_9(=\sigma^2_U)$	11.877	12.324	12.455	13.466	13.490	14.536	14.800	15.397	13.486	13.456	0.297
$\sigma^2_{10}(=\sigma^2_S)$	6.939	7.143	7.246	7.809	7.818	8.447	8.556	9.003	7.811	7.829	0.283
$\sigma^2_{11}(=\sigma^2_H)$	10.480	10.908	10.998	11.710	11.766	12.608	12.818	13.119	11.655	11.803	0.304
$\sigma^2_{12}(=\sigma^2_M)$	3.638	3.779	3.806	4.078	4.092	4.383	4.492	4.614	4.072	4.083	0.292

Table B.19: MCMC results for the marginal parameters of the reduced model of the 3rd vine construction with credible level 10%.

θ	$\hat{\theta}_{min}$	$\hat{\theta}_{2.5\%}$	$\hat{\theta}_{5\%}$	$\hat{\theta}_{med}$	$\bar{\theta}$	$\hat{\theta}_{95\%}$	$\hat{\theta}_{97.5\%}$	$\hat{\theta}_{max}$	$\hat{\theta}_{mod}$	$\hat{\theta}_{IFM}$	$\bar{p}_{acc}$
$\beta_1(=\beta_B)$	1.239	1.258	1.261	1.292	1.292	1.322	1.328	1.340	1.292	1.294	1.000
$\beta_2(=\beta_\$)$	1.119	1.128	1.134	1.160	1.159	1.185	1.193	1.208	1.160	1.158	1.000
$\beta_3(=\beta_N)$	0.728	0.744	0.748	0.768	0.768	0.789	0.794	0.806	0.767	0.767	1.000
$\beta_4(=\beta_E)$	0.797	0.818	0.821	0.858	0.859	0.893	0.898	0.913	0.856	0.858	1.000
$\beta_5(=\beta_D)$	1.171	1.180	1.186	1.216	1.217	1.250	1.255	1.273	1.215	1.218	1.000
$\beta_6(=\beta_T)$	0.590	0.611	0.614	0.646	0.645	0.673	0.680	0.707	0.647	0.643	1.000
$\beta_7(=\beta_O)$	1.086	1.106	1.111	1.135	1.135	1.158	1.162	1.191	1.137	1.135	1.000
$\beta_8(=\beta_C)$	0.935	0.949	0.955	0.980	0.980	1.005	1.007	1.015	0.979	0.978	1.000
$\beta_9(=\beta_U)$	0.748	0.764	0.769	0.806	0.806	0.842	0.853	0.866	0.806	0.805	1.000
$\beta_{10}(=\beta_S)$	0.913	0.929	0.935	0.961	0.961	0.991	0.994	1.015	0.961	0.961	1.000
$\beta_{11}(=\beta_H)$	0.805	0.827	0.835	0.867	0.867	0.900	0.904	0.931	0.866	0.865	1.000
$\beta_{12}(=\beta_M)$	1.147	1.169	1.172	1.188	1.189	1.208	1.213	1.223	1.187	1.189	1.000
$\sigma^2_1(=\sigma^2_B)$	9.519	9.853	9.952	10.614	10.679	11.563	11.670	12.104	10.575	10.521	0.287
$\sigma^2_2(=\sigma^2_\$)$	6.788	6.982	7.089	7.630	7.653	8.202	8.323	8.748	7.625	7.527	0.301
$\sigma^2_3(=\sigma^2_N)$	4.452	4.631	4.668	5.021	5.018	5.410	5.487	5.782	5.023	5.017	0.284
$\sigma^2_4(=\sigma^2_E)$	12.676	13.261	13.453	14.461	14.492	15.715	15.856	16.547	14.473	14.412	0.280
$\sigma^2_5(=\sigma^2_D)$	13.213	13.413	13.571	14.546	14.568	15.673	15.783	16.276	14.514	14.427	0.301
$\sigma^2_6(=\sigma^2_T)$	7.771	8.137	8.301	8.897	8.941	9.631	9.774	10.116	8.844	8.942	0.300
$\sigma^2_7(=\sigma^2_O)$	6.107	6.212	6.324	6.797	6.808	7.291	7.387	7.655	6.790	6.790	0.302
$\sigma^2_8(=\sigma^2_C)$	5.820	6.013	6.088	6.580	6.572	7.064	7.152	7.335	6.575	6.548	0.294
$\sigma^2_9(=\sigma^2_U)$	12.017	12.494	12.575	13.473	13.499	14.385	14.664	15.082	13.438	13.456	0.300
$\sigma^2_{10}(=\sigma^2_S)$	6.760	7.117	7.234	7.789	7.795	8.412	8.514	9.000	7.781	7.829	0.292
$\sigma^2_{11}(=\sigma^2_H)$	10.394	10.813	10.964	11.795	11.821	12.734	12.872	13.288	11.771	11.803	0.291
$\sigma^2_{12}(=\sigma^2_M)$	3.498	3.720	3.791	4.027	4.052	4.353	4.406	4.597	4.005	4.083	0.300

Table B.20: MCMC results for the marginal parameters of the reduced model of the 3rd vine construction with credible level 5%.

θ	$\hat{\theta}_{min}$	$\hat{\theta}_{2.5\%}$	$\hat{\theta}_{5\%}$	$\hat{\theta}_{med}$	$\bar{\theta}$	$\hat{\theta}_{95\%}$	$\hat{\theta}_{97.5\%}$	$\hat{\theta}_{max}$	$\hat{\theta}_{mod}$	$\hat{\theta}_{IFM}$	$C_{10\%}$	$C_{5\%}$	$\bar{p}_{acc}$
$\rho_{1,2}$	-0.312	-0.274	-0.269	-0.220	-0.220	-0.167	-0.159	-0.134	-0.219	-0.224	1	1	0.49
$\rho_{2,3}$	0.130	0.143	0.152	0.205	0.204	0.252	0.260	0.291	0.205	0.212	1	1	0.53
$\rho_{3,4}$	-0.268	-0.236	-0.230	-0.172	-0.173	-0.122	-0.107	-0.083	-0.173	-0.187	1	1	0.50
$\rho_{4,5}$	-0.242	-0.227	-0.221	-0.174	-0.172	-0.127	-0.116	-0.079	-0.174	-0.174	1	1	0.52
$\rho_{5,6}$	-0.213	-0.182	-0.171	-0.122	-0.123	-0.075	-0.068	-0.054	-0.120	-0.120	1	1	0.53
$\rho_{6,7}$	-0.219	-0.177	-0.168	-0.118	-0.118	-0.065	-0.061	-0.029	-0.117	-0.115	1	1	0.54
$\rho_{7,8}$	-0.248	-0.206	-0.197	-0.146	-0.145	-0.088	-0.080	-0.066	-0.149	-0.147	1	1	0.53
$\rho_{9,10}$	-0.156	-0.127	-0.120	-0.062	-0.065	-0.016	-0.005	0.022	-0.062	-0.055	1	1	0.52
$\rho_{10,11}$	0.014	0.039	0.044	0.109	0.107	0.163	0.168	0.210	0.112	0.114	1	1	0.54
$\rho_{11,12}$	-0.223	-0.206	-0.197	-0.147	-0.148	-0.102	-0.094	-0.044	-0.143	-0.153	1	1	0.51
$\rho_{1,3\|\cdot}$	-0.361	-0.340	-0.330	-0.289	-0.287	-0.240	-0.233	-0.198	-0.288	-0.287	1	1	0.47
$\rho_{2,4\|\cdot}$	-0.182	-0.132	-0.116	-0.071	-0.069	-0.017	-0.007	0.008	-0.073	-0.067	1	1	0.52
$\rho_{3,5\|\cdot}$	-0.155	-0.124	-0.115	-0.061	-0.061	-0.006	0.001	0.038	-0.061	-0.065	1	0	0.52
$\rho_{4,6\|\cdot}$	-0.323	-0.280	-0.270	-0.227	-0.228	-0.184	-0.178	-0.135	-0.227	-0.238	1	1	0.52
$\rho_{5,7\|\cdot}$	-0.166	-0.116	-0.109	-0.062	-0.060	-0.009	-0.001	0.035	-0.064	-0.062	1	1	0.54
$\rho_{6,8\|\cdot}$	-0.210	-0.185	-0.177	-0.126	-0.126	-0.074	-0.062	-0.030	-0.127	-0.135	1	1	0.54
$\rho_{7,9\|\cdot}$	-0.152	-0.115	-0.109	-0.059	-0.058	-0.007	0.001	0.028	-0.060	-0.066	1	0	0.52
$\rho_{1,4\|\cdot}$	-0.496	-0.477	-0.472	-0.430	-0.431	-0.393	-0.386	-0.359	-0.430	-0.435	1	1	0.47
$\rho_{4,7\|\cdot}$	-0.235	-0.202	-0.190	-0.139	-0.141	-0.095	-0.089	-0.040	-0.138	-0.133	1	1	0.53
$\rho_{5,8\|\cdot}$	0.101	0.112	0.127	0.171	0.170	0.214	0.225	0.257	0.171	0.178	1	1	0.53
$\rho_{6,9\|\cdot}$	0.032	0.056	0.071	0.121	0.119	0.164	0.171	0.206	0.124	0.125	1	1	0.53
$\rho_{7,10\|\cdot}$	-0.113	-0.083	-0.075	-0.034	-0.033	0.009	0.013	0.046	-0.034	-0.049	0	0	0.50
$\rho_{8,11\|\cdot}$	0.054	0.070	0.079	0.139	0.137	0.192	0.198	0.225	0.141	0.140	1	1	0.53
$\rho_{9,12\|\cdot}$	-0.283	-0.246	-0.235	-0.189	-0.189	-0.145	-0.135	-0.091	-0.191	-0.197	1	1	0.51
$\rho_{1,5\|\cdot}$	-0.187	-0.155	-0.143	-0.085	-0.088	-0.034	-0.027	0.008	-0.084	-0.092	1	1	0.55
$\rho_{4,8\|\cdot}$	-0.143	-0.122	-0.108	-0.065	-0.064	-0.015	-0.005	0.038	-0.066	-0.080	1	1	0.54
$\rho_{5,9\|\cdot}$	-0.145	-0.121	-0.112	-0.065	-0.062	-0.008	0.000	0.016	-0.066	-0.064	1	1	0.54
$\rho_{7,11\|\cdot}$	-0.215	-0.147	-0.140	-0.095	-0.094	-0.044	-0.037	-0.018	-0.096	-0.095	1	1	0.53
$\rho_{8,12\|\cdot}$	0.186	0.213	0.226	0.273	0.273	0.322	0.328	0.351	0.271	0.278	1	1	0.51
$\rho_{1,6\|\cdot}$	-0.255	-0.238	-0.221	-0.175	-0.173	-0.119	-0.114	-0.092	-0.177	-0.181	1	1	0.55
$\rho_{2,7\|\cdot}$	0.018	0.038	0.053	0.106	0.106	0.155	0.162	0.221	0.105	0.106	1	1	0.54
$\rho_{3,8\|\cdot}$	0.072	0.114	0.123	0.165	0.166	0.218	0.225	0.260	0.163	0.175	1	1	0.52
$\rho_{4,9\|\cdot}$	0.000	0.025	0.033	0.083	0.083	0.138	0.147	0.183	0.084	0.079	1	1	0.52
$\rho_{5,10\|\cdot}$	0.055	0.090	0.099	0.148	0.149	0.200	0.207	0.234	0.145	0.144	1	1	0.51
$\rho_{7,12\|\cdot}$	0.209	0.268	0.276	0.323	0.323	0.367	0.375	0.388	0.322	0.326	1	1	0.51
$\rho_{1,7\|\cdot}$	-0.292	-0.273	-0.263	-0.214	-0.214	-0.163	-0.153	-0.097	-0.215	-0.214	1	1	0.55
$\rho_{3,9\|\cdot}$	0.026	0.056	0.065	0.114	0.115	0.166	0.175	0.192	0.112	0.135	1	1	0.52
$\rho_{4,10\|\cdot}$	-0.386	-0.360	-0.355	-0.305	-0.305	-0.263	-0.255	-0.191	-0.306	-0.315	1	1	0.49
$\rho_{5,11\|\cdot}$	-0.301	-0.269	-0.255	-0.208	-0.207	-0.158	-0.151	-0.082	-0.206	-0.214	1	1	0.53
$\rho_{6,12\|\cdot}$	-0.349	-0.324	-0.317	-0.269	-0.268	-0.218	-0.202	-0.147	-0.271	-0.272	1	1	0.53
$\rho_{1,8\|\cdot}$	-0.262	-0.209	-0.200	-0.149	-0.148	-0.095	-0.089	-0.045	-0.150	-0.157	1	1	0.57
$\rho_{2,9\|\cdot}$	0.260	0.306	0.311	0.361	0.360	0.403	0.418	0.463	0.363	0.366	1	1	0.50
$\rho_{3,10\|\cdot}$	0.320	0.360	0.365	0.403	0.404	0.441	0.449	0.486	0.404	0.407	1	1	0.48
$\rho_{4,11\|\cdot}$	-0.244	-0.221	-0.207	-0.158	-0.157	-0.109	-0.098	-0.067	-0.157	-0.168	1	1	0.55
$\rho_{5,12\|\cdot}$	0.062	0.131	0.139	0.197	0.195	0.244	0.265	0.299	0.200	0.197	1	1	0.57
$\rho_{1,9\|\cdot}$	-0.218	-0.170	-0.159	-0.109	-0.110	-0.060	-0.048	-0.017	-0.110	-0.110	1	1	0.57
$\rho_{2,10\|\cdot}$	0.002	0.029	0.038	0.087	0.086	0.138	0.151	0.180	0.088	0.085	1	1	0.57
$\rho_{3,11\|\cdot}$	0.188	0.221	0.229	0.280	0.280	0.326	0.341	0.371	0.278	0.278	1	1	0.54
$\rho_{4,12\|\cdot}$	-0.177	-0.150	-0.140	-0.088	-0.091	-0.043	-0.036	-0.017	-0.085	-0.093	1	1	0.58
$\rho_{1,10\|\cdot}$	-0.184	-0.159	-0.147	-0.102	-0.101	-0.055	-0.046	-0.014	-0.102	-0.102	1	1	0.57

Table B.21: MCMC results for the copula parameters of the reduced model of the 3rd vine construction with credible level 10%.

θ	$\hat{\theta}_{min}$	$\hat{\theta}_{2.5\%}$	$\hat{\theta}_{5\%}$	$\hat{\theta}_{med}$	$\bar{\theta}$	$\hat{\theta}_{95\%}$	$\hat{\theta}_{97.5\%}$	$\hat{\theta}_{max}$	$\hat{\theta}_{mod}$	$\hat{\theta}_{IFM}$	$C_{10\%}$	$C_{5\%}$	$\bar{p}_{acc}$
$\rho_{1,2}$	-0.300	-0.276	-0.268	-0.223	-0.225	-0.183	-0.176	-0.142	-0.221	-0.224	1	1	0.48
$\rho_{2,3}$	0.133	0.158	0.170	0.213	0.215	0.258	0.271	0.290	0.215	0.212	1	1	0.50
$\rho_{3,4}$	-0.263	-0.236	-0.229	-0.178	-0.178	-0.127	-0.116	-0.104	-0.176	-0.187	1	1	0.49
$\rho_{4,5}$	-0.263	-0.230	-0.218	-0.167	-0.168	-0.117	-0.108	-0.101	-0.166	-0.174	1	1	0.52
$\rho_{5,6}$	-0.191	-0.178	-0.167	-0.120	-0.120	-0.071	-0.062	-0.019	-0.123	-0.120	1	1	0.53
$\rho_{6,7}$	-0.193	-0.167	-0.161	-0.114	-0.112	-0.064	-0.052	-0.033	-0.116	-0.115	1	1	0.53
$\rho_{7,8}$	-0.227	-0.200	-0.192	-0.150	-0.149	-0.098	-0.088	-0.067	-0.150	-0.147	1	1	0.53
$\rho_{10,11}$	-0.011	0.037	0.045	0.102	0.102	0.152	0.173	0.205	0.105	0.114	1	1	0.54
$\rho_{11,12}$	-0.249	-0.212	-0.201	-0.153	-0.152	-0.102	-0.092	-0.074	-0.153	-0.153	1	1	0.51
$\rho_{1,3\|\cdot}$	-0.368	-0.338	-0.333	-0.289	-0.289	-0.240	-0.232	-0.190	-0.290	-0.287	1	1	0.48
$\rho_{2,4\|\cdot}$	-0.169	-0.142	-0.131	-0.076	-0.077	-0.027	-0.022	0.008	-0.074	-0.067	1	1	0.52
$\rho_{3,5\|\cdot}$	-0.135	-0.116	-0.103	-0.055	-0.055	-0.009	-0.001	0.014	-0.053	-0.065	1	1	0.53
$\rho_{4,6\|\cdot}$	-0.290	-0.281	-0.273	-0.228	-0.228	-0.182	-0.174	-0.148	-0.226	-0.238	1	1	0.51
$\rho_{5,7\|\cdot}$	-0.156	-0.115	-0.108	-0.058	-0.057	-0.007	0.002	0.022	-0.059	-0.062	1	0	0.54
$\rho_{6,8\|\cdot}$	-0.211	-0.183	-0.178	-0.124	-0.125	-0.076	-0.065	-0.045	-0.123	-0.135	1	1	0.53
$\rho_{1,4\|\cdot}$	-0.508	-0.474	-0.470	-0.424	-0.425	-0.385	-0.376	-0.334	-0.422	-0.435	1	1	0.47
$\rho_{4,7\|\cdot}$	-0.222	-0.203	-0.196	-0.148	-0.148	-0.097	-0.089	-0.064	-0.148	-0.133	1	1	0.52
$\rho_{5,8\|\cdot}$	0.105	0.124	0.128	0.176	0.175	0.218	0.225	0.242	0.178	0.178	1	1	0.53
$\rho_{6,9\|\cdot}$	0.022	0.061	0.073	0.121	0.120	0.168	0.174	0.203	0.120	0.125	1	1	0.53
$\rho_{8,11\|\cdot}$	0.031	0.071	0.084	0.138	0.136	0.182	0.188	0.221	0.141	0.140	1	1	0.52
$\rho_{9,12\|\cdot}$	-0.234	-0.219	-0.209	-0.161	-0.160	-0.111	-0.107	-0.082	-0.159	-0.197	1	1	0.52
$\rho_{1,5\|\cdot}$	-0.171	-0.154	-0.147	-0.095	-0.096	-0.047	-0.038	-0.005	-0.090	-0.092	1	1	0.56
$\rho_{4,8\|\cdot}$	-0.160	-0.126	-0.118	-0.067	-0.067	-0.020	-0.010	0.034	-0.065	-0.080	1	1	0.53
$\rho_{7,11\|\cdot}$	-0.207	-0.161	-0.154	-0.101	-0.102	-0.053	-0.047	-0.027	-0.099	-0.095	1	1	0.53
$\rho_{8,12\|\cdot}$	0.158	0.215	0.221	0.270	0.272	0.323	0.328	0.349	0.270	0.278	1	1	0.50
$\rho_{1,6\|\cdot}$	-0.258	-0.233	-0.226	-0.173	-0.173	-0.125	-0.111	-0.080	-0.170	-0.181	1	1	0.55
$\rho_{2,7\|\cdot}$	0.048	0.073	0.083	0.130	0.129	0.175	0.182	0.209	0.133	0.106	1	1	0.53
$\rho_{3,8\|\cdot}$	0.076	0.104	0.114	0.162	0.161	0.211	0.218	0.240	0.162	0.175	1	1	0.52
$\rho_{4,9\|\cdot}$	-0.009	0.009	0.020	0.072	0.068	0.115	0.124	0.153	0.072	0.079	1	1	0.51
$\rho_{5,10\|\cdot}$	0.063	0.086	0.098	0.151	0.150	0.206	0.217	0.233	0.152	0.144	1	1	0.50
$\rho_{7,12\|\cdot}$	0.247	0.267	0.278	0.325	0.325	0.368	0.375	0.399	0.326	0.326	1	1	0.51
$\rho_{1,7\|\cdot}$	-0.294	-0.273	-0.265	-0.219	-0.220	-0.172	-0.163	-0.123	-0.217	-0.214	1	1	0.55
$\rho_{3,9\|\cdot}$	0.060	0.079	0.095	0.148	0.144	0.189	0.196	0.248	0.151	0.135	1	1	0.50
$\rho_{4,10\|\cdot}$	-0.380	-0.358	-0.354	-0.307	-0.308	-0.262	-0.253	-0.225	-0.308	-0.315	1	1	0.48
$\rho_{5,11\|\cdot}$	-0.288	-0.266	-0.260	-0.212	-0.212	-0.167	-0.155	-0.110	-0.211	-0.214	1	1	0.53
$\rho_{6,12\|\cdot}$	-0.379	-0.332	-0.320	-0.271	-0.272	-0.226	-0.218	-0.186	-0.272	-0.272	1	1	0.53
$\rho_{1,8\|\cdot}$	-0.264	-0.208	-0.200	-0.153	-0.153	-0.101	-0.093	-0.060	-0.156	-0.157	1	1	0.56
$\rho_{2,9\|\cdot}$	0.295	0.309	0.320	0.366	0.366	0.411	0.421	0.438	0.366	0.366	1	1	0.50
$\rho_{3,10\|\cdot}$	0.324	0.350	0.356	0.405	0.403	0.451	0.459	0.468	0.407	0.407	1	1	0.48
$\rho_{4,11\|\cdot}$	-0.243	-0.217	-0.209	-0.161	-0.157	-0.103	-0.097	-0.072	-0.163	-0.168	1	1	0.55
$\rho_{5,12\|\cdot}$	0.116	0.132	0.146	0.196	0.197	0.243	0.253	0.268	0.202	0.197	1	1	0.55
$\rho_{1,9\|\cdot}$	-0.188	-0.168	-0.162	-0.111	-0.112	-0.060	-0.049	-0.025	-0.114	-0.110	1	1	0.57
$\rho_{2,10\|\cdot}$	-0.014	0.023	0.037	0.086	0.086	0.136	0.151	0.194	0.088	0.085	1	1	0.58
$\rho_{3,11\|\cdot}$	0.189	0.222	0.234	0.281	0.282	0.326	0.336	0.361	0.280	0.278	1	1	0.55
$\rho_{4,12\|\cdot}$	-0.164	-0.141	-0.135	-0.090	-0.087	-0.031	-0.017	0.002	-0.091	-0.093	1	1	0.56
$\rho_{1,10\|\cdot}$	-0.203	-0.164	-0.155	-0.103	-0.103	-0.052	-0.042	-0.005	-0.103	-0.102	1	1	0.57

Table B.22: MCMC results for the copula parameters of the reduced model of the 3rd vine construction with credible level 5%.

Bibliography

Aas, K., C. Czado, A. Frigessi, and H. Bakken (2009). Pair-copula construtions of multiple dependence. *Insurance: Mathematics and Economics 44*(2), 182–198.

Baba, K., R. Shibata, and M. Sibuya (2004). Partial correlation and conditional correlation as measures of conditional independence. *Australian & New Zealand Journal of Statistics 46*(4), 657–664.

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

Bickel, P. J. and K. A. Doksum (2001). *Mathematical Statistics: Basic Ideas and Selected Topics (2nd ed.)*. Prentice Hall.

Congdon, P. (2003). *Applied Bayesian Modelling*. Wiley & Sons.

Czado, C. and A. Min (2009). Bayesian inference for multivariate copulas using pair-copula constructions. *Working Paper*.

Gelfand, A. E. and A. F. M. Smith (1990). Sampling based approaches to calculating marginal densities. *Journal of the American Statititstical Association 85*, 398–409.

Geman, S. and D. Geman (1984). Stochastic relaxation, gibbs distributions, and the bayesian restoration of images. *IEEE Transaction Pattern Analysis and Machine Intelligence 6*.

Georgii, H.-O. (2002). *Einführung in die Wahrscheinlichkeitstheorie und Statistik*. de Gruyter.

Gilks, W., S. Richardson, and D. Spiegelhalter (1996). Introducing markov chain monte carlo. In W. Gilks, S. Richardson, and D. Spiegelhalter (Eds.), *Markov Chain Monte Carlo in Practice*, pp. 1–19. Chapman & Hall.

Gschößl, S. (2006). *Hierarchical Bayesian Spatial Regression Models with Applications in Non-Life Insurance*. Ph. D. thesis, Mathematische Fakultät, Technische Universität München.

Hastings, W. K. (1970). Monte carlo sampling methods using markov chains and their applications. *Biometrika 57*.

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*.

Joe, H. and J. J. Xu (1996). The estimation method of inference functions for margins for multivariate models. Technical report, Department of Statistics, University of British Columbia.

Kastenmeier, R. (2008). Joint regression analysis of insurance claims and claim size. Master's thesis, Mathematische Fakultät, Technische Universität München.

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*.

Kurowicka, D. and R. M. Cooke (2006). *Uncertainty Analysis with High Dimensional Dependence Modelling*. Wiley & Sons.

Lewandowski, D., D. Kurowicka, and H. Joe (2007). Generating random correlation matrices based on vines and extended onion method. *submitted to Journal of Multivariate Analysis*.

Lintner, J. (1965). The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. *The Review of Economics and Statistics 47*(1), 13–37.

Markowitz, H. (1952). Portfolio selection. *The Journal of Finance 7*(1), 77–91.

Metropolis, N., A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller (1953). Equation of state calculations by fast computing machines. *The Journal of Chemical Physics 21*(6), 1087–1092.

Nelsen, R. B. (1999). *An Introduction to Copulas* (2nd ed.). Springer Series in Statistics. New York: Springer.

Pitt, M., D. Chan, and R. Kohn (2006). Efficient bayesian inference for gaussian copula regression models. *Biometrika 93*(3), 537–554.

Seeley, R. T. (1961). Fubini implies Leibniz implies $F_{yx} = F_{xy}$. *The Journal of Chemical Physics 68*, 56–57.

Sharpe, W. F. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. *The Journal of Finance 19*(3), 425–442.

Silva, R. d. S. and H. F. Lopes (2008). Copula, marginal distributions and model selection: a bayesian note. *Statistics and Computing 18*(3), 313–320.

Sklar, A. (1959). Fonctions de répartition à n dimensions et leurs marges. *Publications de l'Institut de Statistique de l'Université de Paris 8*, 229–231.

Spiegelhalter, S. D., N. G. Best, B. P. Carlin, and A. V. D. Linde (2002). Bayesian measures of model complexity and fit (with discussion). *Journal of the Royal Statistical Society: Series B 64*(4), 583–639.

Venables, W. N. and B. D. Ripley (2003). *Modern Applied Statistics with S*. Springer.

Yule, G. U. and M. G. Kendall (1965). *An Introduction to the Theory of Statistics*. Springer Series in Statistics. Charles Griffin & Company.