Bayesian Risk Aggregation:
Correlation Uncertainty and Expert Judgement

Klaus Böcker, Alessandra Crimmi and Holger Fink

Presented at
2010 Enterprise Risk Management Symposium
Society of Actuaries
April 12-15, 2010

Bayesian Risk Aggregation: Correlation Uncertainty and Expert Judgement

Klaus Böcker * Alessandra Crimmi † Holger Fink ‡

1 Introduction

In this Chapter we present a novel way for estimating aggregated EC figures based on Bayesian copula estimation. Contrary to the classic approach of using a single inter-risk-correlation matrix we derive a probability distribution of possible correlation matrices that enables us to tackle the important issue of *parameter uncertainty*.

One of the main concerns in connection with risk aggregation is of whether and to which extent diversification benefits between different risk types can be identified. Apart from the very simple approach of adding-up all EC estimates for each risk category or business line one can distinguish between *top-down* or modular approaches on the one hand and *bottom-up* or multi-factor simulation approaches on the other, cf. for instance Saita [24] for further information and references about principles of risk aggregation.

Bottom-up approaches basically model all the bank's real and financial variables including assets, liabilities and interest sensitive off-balance sheet items simultaneously, see for e.g. Kretzschmar, McNeil, and Kirchner [16]. This allows to capture gains and losses at the level of individual instruments or positions without the need for creating artificial risk silos. Such sophisticated approaches are particularly important for market and credit risk, which are highly related and inextricably linked with each other, see e.g. the Research Task Force of the Basel Committee on Banking Supervision [2] and Hartmann, Pritsker, and Schuermann [14] for more details about the interaction of market and credit risk.

While financial institutions and supervisors are seeking for flexible bottom-up methods for the aggregation of market and credit risk, conceptual difficulties remain with respect to

*Risk Analytics and Methods, UniCredit Bank AG, München, Germany,
email: klaus.boecker@unicreditgroup.de

†Risk Analytics and Methods, UniCredit Group, Milan, Italy,
email: alessandra.crimmi@unicreditgroup.eu

‡Center for Mathematical Sciences, Technische Universität München, Garching bei München,
Germany, email: fink@ma.tum.de

other risk types such as operational risk and business risk. As a matter of fact, banks are currently still favoring simpler top-down methods when computing their aggregated EC as pointed out by the IFRI/CRO Forum [15]. According to this survey the most popular method in practice is the *aggregation-by-risk-type* approach where stand-alone risk figures of different risk types are combined in some way to obtain the desired aggregated EC. In a similar, more recent survey of the Basel Committee [3] it is reported that "*there is no established set of best practices concerning risk aggregation in the industry.*" From all this, it can be expected that for quite some time hybrid approaches that at least partially rely on an inter-risk-correlation matrix will heavily influence market practices.

The simplest form of risk aggregation expresses the dependence between different risk types by an *inter-risk-correlation matrix* $\boldsymbol{R}$, and its estimation and calibration is a core problem for the calculation of total EC in practice. A standard approach is to model the dependence structure between risk types by a distributional copula, see e.g. the references in Böcker [5]. Estimates for inter-risk-correlations differ significantly within the industry. The IFRI/CRO Forum [15] points out that "*correlation estimates used vary widely, to an extent that is unlikely to be solely attributable to differences in business mix.*" More bluntly, one could also say that banks often have only a vague opinion about inter-risk correlations and that their estimation is afflicted with high uncertainties. One reason for this is that very often reliable data are scarce and do not cover long historical time periods. Therefore, inter-risk correlations are approximated by the co-movement of asset price indices or similar proxies of which it is hoped are representative for these risk types. As a consequence thereof, a reliable and robust statistical estimate of the inter-risk-correlation matrix is often not possible and it is necessary to draw on expert opinions. This has recently also been acknowledged by the Committee of European Banking Supervisors [8], where they explicitly distinguish statistical techniques versus expert judgements.

Our work makes two novel contributions for estimating aggregated EC within an aggregation-by-risk-type framework. First, we explicitly address the existence of parameter uncertainty associated with the inter-risk-correlation matrix. Second, we present a sophisticated method for assessing inter-risk correlations (more precisely, the Gaussian copula parameters) by means of expert judgement. To illustrate our approach, we calculate aggregated EC for the same portfolio already used in Böcker [5], consisting of 10 % market risk, 61 % credit risk, 14 % operational risk, and 15 % business risk in terms of 99.95 % EC. Here, however, we make an assumption which is key to what follows, namely that as a consequence of all the uncertainties *a bank's inter-risk-correlation matrix cannot be considered as a fixed parameter of the risk-aggregation model but should rather be treated as a random parameter.* Hence, in a Bayesian framework, the inter-risk-correlation matrix $\boldsymbol{R}$ (or, more generally, the parameters of the copula) is described by a *posterior distribu-*

tion, which comprises empirical information (e.g. time series of risk proxies representative for each risk type) as well as expert judgement.

This Chapter is organised as follows. First we briefly recap the "classic" aggregation-by-risk-type approach using a fixed Gaussian copula and also introduce the portfolio that serves as an illustrative example. Then we describe the construction of the prior and posterior distribution of the inter-risk-correlation matrix by considering the different pair correlations separately. For specific models of the pair correlation priors (the beta model, the triangular model, and the uniform model) we suggest a Markov-Chain-Monte-Carlo (MCMC) algorithm that can easily be used to sample a set of correlation matrices from the posterior. We then discuss a numerical example of risk aggregation and, finally, we examine how inter-risk correlations may be estimated using expert knowledge.

2 Classic copula aggregation

A d-dimensional distributional copula C is a d-dimensional distribution function on $[0,1]^d$ with uniform marginals. Among all copulas discussed in the literature, maybe those most frequently used for risk aggregation are the Gaussian copula and the t copula. The importance of copulas for financial risk management is essentially a result of Sklar's theorem, stating that every multivariate distribution function can be separated into their marginal distribution functions and a copula. Therefore, copulas allow for a separate modelling of the marginals of distinct risk types on the one hand and their dependence structure (i.e. the copula) on the other. Distributional copulas have been frequently applied to risk aggregation e.g. Dimakos & Aas [9], Rosenberg & Schuermann [23], Ward & Lee [26], or Brockmann & Kalkbrener [6].

In the sequel, we summarise some properties of the Gaussian copula, see Cherubini, Luciano, and Vecchiato [7] for more details. Let Φ denote the standard univariate normal distribution function, $\Phi^d_{\boldsymbol{R}}$ the standard multivariate normal distribution function with $d \times d$ correlation matrix $\boldsymbol{R}$, which equals the covariance matrix, and $(u_1, \dots, u_d) \in [0,1]^d$. Then, the distribution function of the d-dimensional Gaussian copula is given by

$$C^d_{\boldsymbol{R}}(u_1, \dots, u_d) = \Phi^d_{\boldsymbol{R}}(\Phi^{-1}(u_1), \dots, \Phi^{-1}(u_d)) \tag{2.1}$$

where $\Phi^{-1}(\cdot)$ denotes the inverse of the standard normal distribution function. The density of the d-dimensional Gaussian copula can be written as

$$c^d_{\boldsymbol{R}}(u_1, \dots, u_d) = \det(\boldsymbol{R})^{-\frac{1}{2}} \exp\left[-\frac{1}{2}\boldsymbol{\xi}'(\boldsymbol{R}^{-1} - \boldsymbol{I}_d)\boldsymbol{\xi}\right] \tag{2.2}$$

with $\boldsymbol{\xi} = (\Phi^{-1}(u_1), \dots, \Phi^{-1}(u_d))'$ and identity matrix $\boldsymbol{I}_d$.

We now turn to the estimation of the Gaussian copula which is usually done by maximising the likelihood function. Suppose $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$ is an n-sample of $d \times 1$ mutually independent observations that are identically distributed. In the context of inter-risk aggregation each marginal component of the vector $\boldsymbol{x}_j, j = 1, \ldots, n$, represents a suitable risk driver or loss proxy representative for a different risk type. Specifically, assuming a Gaussian copula model means that all components of $\boldsymbol{x}_j, j = 1, \ldots, n$, have a Gaussian dependence structure.

After transforming the sample data $\boldsymbol{x}_j$ into variates $\boldsymbol{u}_j$ with uniform marginals (e.g. using order statistics or parametric distribution functions), we obtain the likelihood of the Gaussian copula as (sse Press [20])

$$l(\boldsymbol{R}|\boldsymbol{\xi}_1, \ldots, \boldsymbol{\xi}_n) \propto \det(\boldsymbol{R})^{-\frac{n}{2}} \exp\left[-\frac{n}{2}\mathrm{tr}(\boldsymbol{R}^{-1}\boldsymbol{B})\right] , \tag{2.3}$$

where the $d \times d$ symmetric, positive semidefinite matrix $\boldsymbol{B}$ denotes the sample covariance matrix of the data after transformation to standard normal marginals, i.e.

$$\boldsymbol{B} = \frac{1}{n}\sum_{j=1}^{n} \boldsymbol{\xi}_j \boldsymbol{\xi}_j' \quad \text{with} \quad \boldsymbol{\xi}_j = (\Phi^{-1}(u_{1j}), \ldots, \Phi^{-1}(u_{dj}))' . \tag{2.4}$$

The matrix $\boldsymbol{B}$ is also the global maximum of the likelihood function (2.3) (see e.g. Press [20], p. 183) from which the Gaussian copula parameter $\hat{\boldsymbol{R}}$ can be obtained.

After the matrix $\hat{\boldsymbol{R}}$ has been determined, one can start with the risk-type aggregation. As mentioned in the introduction, we adopt the portfolio used in Böcker [5] consisting of 10 % market risk, 61 % credit risk, 14 % operational risk, and 15 % business risk, representing an industry average obtained from large banks reported in IFRI/CRO Forum [15], Figure 16. Furthermore, we assume that EC for each risk type was calculated at confidence level of 99.95 %, time horizon of one year, and that the marginal distribution functions are as in Table 2.1. Now, observing that the inverses $F_i^{-1}(\cdot), i =, 1, \ldots, d$, of all marginal risk type distribution functions $F_i(\cdot)$ exist, classic risk aggregation with a Gaussian copula is straight-forward. In a first step, one simulates a large number N of mutually independent $d \times 1$ vectors $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_N$ from the Gaussian copula $C^d_{\hat{\boldsymbol{R}}}$. Note that by definition for all $j = 1, \ldots, N$ the marginal components $\{u_{1j}, \ldots, u_{dj}\}$ of the vectors $\boldsymbol{u}_j$ are uniformly distributed with a Gaussian dependence structure. In a second step, all marginal components are transformed by $F_i^{-1}(u_{ij}),\ i = 1, \ldots, d,\ j = 1, \ldots, N$. Since copulas are invariant under strictly increasing transformations, the dependence structure between $\{F_1^{-1}(u_{1j}), \ldots, F_d^{-1}(u_{dj})\}_{j=1,\ldots,N}$ is the same as between $\{u_{1j}, \ldots, u_{dj}\}_{j=1,\ldots,N}$, that is a Gaussian one. Finally, a simulated N-sample of aggregated EC can be computed through $\left\{ \sum_i^d F_i^{-1}(u_{ij})\right\}_{j=1,\ldots,N}$.

Risk	Distribution function	Parameters
MR	$F(x) = F_\nu\left(\frac{x-\mu}{\sigma}\right), \quad x \in \mathbb{R}$	$\mu = 0,\ \sigma = 2.18, \nu = 10$
CR	$F(x) = \Phi\left[\frac{1}{\sqrt{\varrho}}\left(\sqrt{1-\varrho}\,\Phi^{-1}(\frac{x}{X}) - \Phi^{-1}(p)\right)\right], \quad x > 0$	$X = 2338.64$, $p = 0.3\%, \varrho = 8\%$
OR	$F(x) = \Phi\left[\frac{\ln x - \mu}{\sigma}\right], \quad x > 0$	$\mu = -0.893, \sigma = 1.089$
BR	$F(x) = \Phi\left[\frac{x-\mu}{\sigma}\right], \quad x \in \mathbb{R}$	$\mu = 0,\ \sigma = 4.56$

Table 2.1: Marginal distributions for market risk (MR), credit risk (CR), operational risk (OR), and business risk (BR), where F_ν is the Student-t distribution function with ν degrees of freedom and Φ is the standard normal distribution function. MR follows a scaled Student-t and CR is described by a Vasicek distribution with total exposure X, uniform asset correlation ϱ, and average default probability p. Operational risk (OR) is assumed to be lognormally distributed and business risk (BR) is modelled by a normal distribution. The parameters for the distribution functions are chosen so that MR, CR, OR, and BR absorb 10, 61, 14, and 15 units of EC at 99.95 % confidence level. Finally, only credit risk and operational risk have non-zero expected losses of about 7.0 and 0.7, respectively.

3 Bayesian risk aggregation

3.1 Construction of the inter-risk-correlation prior

In the Bayesian approach for risk-type aggregation, one has to find a suitable prior for the inter-risk-correlation matrix $\boldsymbol{R}$ of the Gaussian copula. This prior represents the available expert knowledge regarding the inter-risk-correlation matrix in probabilistic form. In our example the 4×4 correlation matrices $\boldsymbol{R}$ can be identified with a 6-dimensional real vector by

$$\boldsymbol{R} = \begin{pmatrix} 1 & r_1 & r_2 & r_3 \\ \cdot & 1 & r_4 & r_5 \\ \cdot & \cdot & 1 & r_6 \\ \cdot & \cdot & \cdot & 1 \end{pmatrix} \longleftrightarrow (r_1, r_2, r_3, r_4, r_5, r_6), \qquad r_i \in [-1, 1],\ 1 \le i \le 6, \tag{3.1}$$

where the matrix $\boldsymbol{R}$ must be positive semidefinite.

The most commonly used prior model for a *covariance* matrix $\boldsymbol{\Sigma}$ is the inverse-Wishart distribution, see e.g. Press [20]. Since every covariance matrix $\boldsymbol{\Sigma}$ is related to a correlation matrix $\boldsymbol{R}$ by

$$\boldsymbol{\Sigma} = \boldsymbol{S}^{1/2}\boldsymbol{R}\,\boldsymbol{S}^{1/2}\,, \tag{3.2}$$

where $\boldsymbol{S} = \operatorname{diag}(\sigma_{11}, \sigma_{22}, \ldots)$ and σ_{jj} are the diagonal elements of $\boldsymbol{\Sigma}$, the inverse-Wishart prior can also be used to construct a prior for the correlation matrix. An example of this method can be found in the Chapter of Dalla Valle in this volume.

The inverse-Wishart-based prior for the inter-risk-correlation matrix allows only for one single degrees of freedom parameter ν to express prior beliefs (and thus the level of uncertainty) about $\boldsymbol{R}$. In practice, however, the amount of available expert information for different pair correlations $r_i, i = 1, \ldots, 6$, may significantly depend on the risk-type combinations. Another drawback of the inverse-Wishart distribution is that the resulting prior for $\boldsymbol{R}$ cannot easily be estimated by means of expert judgement. Note, that the prior distribution for $\boldsymbol{R}$ is calculated from the inverse-Wishart prior for $\boldsymbol{\Sigma}$ through the transformation (3.2), which rarely yields a prior distribution that can easily be specified by expert elicitation. All this leads to the conclusion that the inverse-Wishart based prior is inadequate for our purpose and that we have to construct a more flexible prior directly for the correlation matrix $\boldsymbol{R}$.

We therefore use a kind of "bottom-up" approach to build the prior distribution for $\boldsymbol{R}$. In a first step, prior information for each component $r_i, i = 1, \ldots, 6$, of $\boldsymbol{R}$ is separately modelled by one-dimensional distributions with Lebesgue densities $\pi_i(r_i), i = 1, \ldots, 6$. By (3.1) this yields a distribution of symmetric, real-valued matrices with diagonal elements equal to one. In a second step we restrict to those matrices preserving the positive semidefiniteness of the correlation matrix. Hence, denoting the space of all 4-dimensional correlation matrices by $\mathfrak{R}^4$, a possible density for the correlation matrix prior can be written as

$$\pi(\boldsymbol{R}) = \prod_{1 \leq i \leq 6} \pi_i(r_i)\, \mathbb{1}_{\{\boldsymbol{R} \in \mathfrak{R}^4\}}. \tag{3.3}$$

The indicator function $\mathbb{1}_{\{.\}}$ ensures that the matrices are positive definite and also introduces a dependence structure among the r_i for $i = 1, \ldots, 6$.

Pairwise correlation priors As described in more detail in a later section of this Chapter, a well-established approach for the subjective determination of a prior density is by matching a given functional form. The shape of the prior reflects the amount and the quality of the available information and should match the experts' beliefs as closely as possible.

As risk managers are concerned about unreasonable and possibly incorrect diversification benefits, it is normally assumed that correlations between different risk types are non-negative. Such a boundary condition for the correlation matrix can easily and naturally be modelled within the Bayesian framework by considering only pairwise priors π_i that have support in [0,1].

Example 3.1. [Uniform prior]
Assume that the experts are totally uninformed about the possible values of the single pair correlations r_i for $i = 1, \ldots, 6$. Consequently, we may want to take all values of r_i as

equally likely and, in consideration of the general restriction $r_i \in [0,1]$, a natural *diffuse* prior is the uniform distribution

$$\pi_i(r_i) \propto \mathbb{1}_{\{0<r_i<1\}}\,, \qquad i = 1,\ldots,6\,.$$

Note, however, that owing to the positive definiteness constraint of $\boldsymbol{R}$, the marginal priors for the individual correlations r_i resulting from (3.3) are not uniformly distributed anymore. See also Barnard, McCulloch, and Meng [1] for further discussion. □

Example 3.2. [Beta distributed prior]
Suppose the pairwise correlations r_i follow a beta distribution,

$$r_i \sim \mathrm{Be}(\alpha_i, \beta_i)\,, \qquad i = 1,\ldots,6,$$

with hyperparameters $\alpha_i, \beta_i > 0$. This approach was also suggested by Gokhale & Press [13] to model the correlation coefficient in a bivariate normal distribution. The beta density is given by

$$\mathrm{Be}(r_i|\alpha_i,\beta_i) = \frac{r_i^{\alpha_i-1}(1-r_i)^{\beta_i-1}}{B(\alpha,\beta)}\,, \qquad r_i \in [0,1]\,, \tag{3.4}$$

where $B(\cdot,\cdot)$ denotes the Euler beta function. The mean value and variance are given by

$$\begin{aligned} \mu_i &= \frac{\alpha_i}{\alpha_i+\beta_i}\,, \\ \sigma_i^2 &= \frac{\alpha_i\beta_i}{(\alpha_i+\beta_i)^2(1+\alpha_i+\beta_i)}\,, \qquad i = 1,\ldots,6\,, \end{aligned} \tag{3.5}$$

which can be utilized to calculate the hyperparameters α_i, β_i by means of expert judgement (see the last section in this Chapter). □

Example 3.3. [Triangular distributed prior]
An alternative family of useful prior distributions for a correlation parameter is the symmetric triangular distribution giving values between α_i and β_i with $-1 \le \alpha_i < \beta_i \le 1$. Then, the prior density for all $r_i, i = 1,\ldots,6$, is of the form

$$T(r_i|\alpha_i,\beta_i) = 4\frac{\beta_i - r_i}{(\beta_i-\alpha_i)^2}\,\mathbb{1}_{\{(\alpha_i+\beta_i)/2<r_i\le\beta_i\}} - 4\frac{\alpha_i - r_i}{(\beta_i-\alpha_i)^2}\,\mathbb{1}_{\{\alpha_i\le r_i\le(\alpha_i+\beta_i)/2\}}, \quad r_i \in \mathbb{R}. \tag{3.6}$$

Similarly to the previous example, the mean and variance can be calculated as

$$\begin{aligned} \mu_i &= \frac{\alpha_i+\beta_i}{2}\,, \\ \sigma_i^2 &= \frac{1}{24}(\alpha_i-\beta_i)^2\,, \qquad i = 1,\ldots,6\,. \end{aligned} \tag{3.7}$$

In contrast to the beta distribution, the support of the triangular distribution is not confined to the interval $[0,1]$. So as to acknowledge the conservative assumption $r_i \in [0,1]$, one can use a truncated version of the triangular distribution instead and the subsequent calculations can be done in a similar way. □

Posterior for the correlation matrix To construct the posterior distribution for the entire inter-risk-correlation matrix of a Gaussian copula, one has to combine the prior distribution (3.3) with the likelihood function (2.3) of the data (after transformation to standard normal marginals) according to Bayes theorem. One then obtains

$$\begin{aligned} p(\boldsymbol{R}|\boldsymbol{\xi}_1,\ldots,\boldsymbol{\xi}_n) &\propto \pi(\boldsymbol{R})\, l(\boldsymbol{R}|\boldsymbol{\xi}_1,\ldots,\boldsymbol{\xi}_n) \\ &\propto \det(\boldsymbol{R})^{-\frac{n}{2}} \exp\left[-\frac{n}{2}\mathrm{tr}(\boldsymbol{R}^{-1}\boldsymbol{B})\right] \prod_{i=1}^{6} \pi_i(r_i)\, \mathbb{1}_{\{\boldsymbol{R}\in\mathfrak{R}^4\}}\,, \end{aligned} \tag{3.8}$$

where $\pi_i(\cdot)$ are the pairwise correlation priors which can, for instance, be chosen according to the examples above (i.e. uniform, beta, or triangular distributed).

3.2 Simulation of inter-risk-correlation matrices

The posterior distribution of the inter-risk-correlation matrix of a Gaussian copula as presented by (3.8) is not a standard distribution. Therefore, we apply MCMC methods to generate a sample of inter-risk-correlation matrices distributed according to $p(\boldsymbol{R}|\boldsymbol{\xi}_1,\ldots,\boldsymbol{\xi}_n)$. MCMC methods entails repeated sampling from a Markov chain that converges to sampling from the posterior distribution, in our case (3.8). A modern overview about computational tools for Bayesian statistics is given by Robert and Rousseau in Chapter 1 in this book; established textbook references on MCMC are Gilks, Richardson, and Spiegelhalter [12], and Robert and Casella [21].

Gibbs sampling One possibility is to simultaneously simulate a 6-dimensional Markov chain of the vector of pair correlations $(r_1,\ldots,r_6)$, e.g. by using a Metropolis-Hastings algorithm. This would necessitate a six-dimensional proposal distribution and our experience is that the convergence of the chain often becomes very slow, especially when more than only four risk types are considered. An alternative and convenient method is Gibbs sampling, which allows to circumvent high dimensionality by simulating componentwise using the *full conditionals* of (3.8) with respect to all but one pair correlation r_i. More precisely, up to a constant, the full conditional posterior distributions for $i=1,\ldots,6$ can be written as

$$p(r_i|r_j, i\neq j) \propto \det\big(\boldsymbol{R}_i(r_i)\big)^{-\frac{n}{2}} \exp\left[-\frac{n}{2}\mathrm{tr}\big(\boldsymbol{R}_i(r_i)^{-1}\boldsymbol{B}\big)\right] \pi_i(r_i)\, \mathbb{1}_{\{\boldsymbol{R}_i(r_i)\in\mathfrak{R}^4\}}\,,\ r_i\in[0,1[\tag{3.9}$$

where $\boldsymbol{R}_i(\cdot) \equiv \boldsymbol{R}(\cdot|r_1,\ldots,r_{i-1},r_{i+1},\ldots,r_6)$ is the correlation matrix obtained from $\boldsymbol{R}$ by fixing all but the i-th pair correlation. These one-dimensional distributions are still complex and not at all standard. However, an independent Metropolis-Hastings algorithm where the proposal density is independent of the current chain value works quite well for our set-up, see details below.

The Gibbs sampler generates an autocorrelated Markov chain of vectors $(r_1^{(t)}, \ldots, r_6^{(t)})_{t=0,1,2,\ldots}$ with stationary distribution $p(\boldsymbol{R}|\boldsymbol{\xi}_1, \ldots, \boldsymbol{\xi}_n)$ given by equation (3.8). The updating of the t-th component of the chain to the $(t+1)$-th component works componentwise by sampling from the one-dimensional full conditionals (3.9):

$$
\begin{aligned}
&(1)\ \ r_1^{(t+1)} \ \sim\ p(r_1|r_2^{(t)}, r_3^{(t)}, \ldots, r_6^{(t)}),\\
&(2)\ \ r_2^{(t+1)} \ \sim\ p(r_2|r_1^{(t+1)}, r_3^{(t)}, \ldots, r_6^{(t)}),\\
&\qquad \vdots\\
&(6)\ \ r_6^{(t+1)} \ \sim\ p(r_6|r_1^{(t+1)}, r_2^{(t+1)}, \ldots, r_5^{(t+1)}).
\end{aligned}
$$

The Gibbs sampler converges in our situation by construction, see Section 10.2 of Robert and Casella [21]. Therefore, after a sufficiently long *burn-in* period of b iterations, the matrices $\boldsymbol{R}^{(t)}_{b<t\leq T}$ built from $(r_1^{(t)}, \ldots, r_6^{(t)})$ are approximately distributed according to the posterior (3.8).

Metropolis-Hastings-within-Gibs sampling The Gibbs algorithm above involves iterated sampling from the full conditional distributions, which in our case is done by an independent Metropolis-Hastings algorithm. Hence, the entire procedure may be referred to as a *Metropolis-Hastings-within-Gibbs* algorithm.

The Metropolis-Hastings algorithm requires an appropriate proposal density. In general, the Metropolis-Hastings algorithm is more efficient when the proposal density is at least approximately similar to the target density, i.e. the full conditionals (3.9). Assuming that the length n of the empirical time series is relatively short, we may suppose that the shape of the full conditional posteriors (3.9) are mainly impacted by the full conditionals of the correlation matrix priors, i.e. by

$$\pi(r_i|r_j, i \neq j) \propto \pi_i(r_i)\, \mathbb{1}_{\{\boldsymbol{R}_i(r_i) \in \mathfrak{R}^4\}}, \quad i = 1, \ldots, 6, \quad r_i \in [0,1]\,. \tag{3.10}$$

Therefore, when deploying the one-dimensional independent Metropolis-Hastings algorithm to sample from (3.9), it may be justified to chose the full conditionals (3.10) as proposal densities. An alternative proposal is the uniform distribution. It should be mentioned that the question regarding which proposal distribution works best can only be answered in the context of the concrete data at hand. In our exercise below, we found that proposals of the form (3.10) are superior to a uniform distribution when beta distributed priors are used (Exercise 3.2). However, in case of triangular shaped priors (Exercise 3.3) either a uniform distribution or the full conditional priors are suitable proposal densities and lead to a good convergence of the MCMC simulation.

When sampling $r_i, i = 1, \ldots, 6$, from (3.9) one has to account for the fact that the resulting matrix $\boldsymbol{R}$ must be positive semidefinite. In order to achieve a maximum computational efficiency, it would be good to know what values of r_i, given all the other correlations $r_j, j \neq i$, keep $\boldsymbol{R}$ positive semidefinite. Following Barnard et al. [1] we remark that the indicator function $\mathbb{1}_{\{\boldsymbol{R}_i(r_i) \in \mathfrak{R}^4\}}$ whose evaluation involves computations of determinants can be rewritten as $\mathbb{1}_{\{r_i \in [a_i(r_j, j \neq i), b_i(r_j, j \neq i)]\}}$ for all $i = 1, \ldots, 6$. Here, $a_i(r_j, j \neq i)$ and $b_i(r_j, j \neq i)$ are the roots of the two-grade polynomial $\det(\boldsymbol{R}_i(r_i)) = 0$ with $\boldsymbol{R}_i(\cdot)$ as defined in (3.9). Therefore, sampling from the full conditionals (3.9) reduces to sampling from the related truncated distributions with truncation intervals $[a_i(\cdot), b_i(\cdot)]$ that can easily be calculated in closed-form[1].

Before we specify the simulation algorithms for different prior assumptions, we should mention that our independent Metropolis-Hastings algorithm always converges due to compact support of proposal and posterior densities, cf. Theorem 7.8 of Robert and Casella [21].

Example 3.4. [Beta distributed prior (continued)]
We use the beta distributed pairwise priors introduced in Example 3.2 and thus the prior for the correlation matrix follows from (3.3) to be

$$\pi(\boldsymbol{R}) = \prod_{1 \leq i \leq 6} \mathrm{Be}(r_i | \alpha_i, \beta_i) \mathbb{1}_{\{\boldsymbol{R} \in \mathfrak{R}^4\}}. \tag{3.11}$$

As already mentioned above, we take as proposal distributions for the independent Metropolis-Hastings algorithm the full conditionals of (3.11), which can be written as

$$\pi(r_i | r_j, j \neq i) \quad \propto \quad \mathrm{Be}(r_i | \alpha_i, \beta_i)\, \mathbb{1}_{\{r_i \in [a_i(r_j, j \neq i), b_i(r_j, j \neq i)]\}}\,, \qquad i = 1, \ldots, 6\,, \tag{3.12}$$

where $a_i(\cdot)$ and $b_i(\cdot)$ are the solutions to $\det(\boldsymbol{R}_i(r_i)) = 0$. We now can specify the whole simulation algorithm as follows:

(I) Choose a starting correlation matrix $\boldsymbol{R}^{(0)} \longleftrightarrow (r_1^{(0)}, \ldots, r_6^{(0)})$ and set $t = 0$.

(II) Set $(z_1, \ldots, z_6) = (r_1^{(t)}, \ldots, r_6^{(t)})$. For $i = 1, \ldots, 6$ do:

(1) Set $k = 0$, $x^{(k)} = z_i$ and define $\boldsymbol{R}^i[\cdot] \equiv \boldsymbol{R}\big[\, \cdot \, | z_j, j \neq i\big]$.

(2) Generate a beta proposal $y \sim \pi(\cdot | z_j, j \neq i)$ according to (3.12).

[1] The brute force method would be to generate samples from the untruncated proposal distribution until the matrix $\boldsymbol{R}$ is checked to be positive semidefinite.

(3) Calculate the update probability δ as

$$\delta = \min\left\{1, \det\left(\frac{\boldsymbol{R}^i[y]}{\boldsymbol{R}^i[x^{(k)}]}\right)^{-\frac{n}{2}} \exp\left[-\frac{n}{2}\mathrm{tr}\big[\big(\boldsymbol{R}^i[y]^{-1} - \boldsymbol{R}^i[x^{(k)}]^{-1}\big)\boldsymbol{B}\big]\right]\right\}.$$

(4) Take $x^{(k+1)} = \begin{cases} y, & \text{with probability } \delta, \\ x^{(k)}, & \text{else.} \end{cases}$

(5) If $k = I_{\mathrm{MH}} - 1$ then stop and set $z_i = x^{(k+1)}$, else set $k = k+1$ and go to (2).

(III) Set $(r_1^{(t+1)}, \ldots, r_6^{(t+1)}) = (z_1, \ldots, z_6)$.
If $t = I_{\mathrm{Gibbs}} - 1$ then stop, else set $t = t+1$ and go to (II).

I_{Gibbs} and I_{MH} determine the number of steps of the Gibbs sampler and the independent Metropolis-Hastings algorithms, respectively. □

Example 3.5. [Triangular distributed prior (continued)]
In case of the triangular priors (3.6) of Example 3.3 we decided to use uniformly distributed proposals. The simulation can be done using the following algorithm:

(I) Choose a starting correlation matrix $\boldsymbol{R}^{(0)} \longleftrightarrow (r_1^{(0)}, \ldots, r_6^{(0)})$ and set $t = 0$.

(II) Set $(z_1, \ldots, z_6) = (r_1^{(t)}, \ldots, r_6^{(t)})$. For $i = 1, \ldots, 6$ do:

(1) Set $k = 0$, $x^{(k)} = z_i$ and define $\boldsymbol{R}^i[\cdot] \equiv \boldsymbol{R}\big[\,\cdot\,| z_j, j \neq i\big]$.

(2) Generate a uniform proposal $y \sim \mathbb{1}_{\{\boldsymbol{R}^i[\cdot] \in \mathfrak{R}^4\}} = \mathbb{1}_{\{\cdot \in [a_i(z_j, j\neq i), b_i(z_j, j \neq i)]\}}$.

(3) Calculate the update probability δ as

$$\delta = \min\left\{1, \det\left(\frac{\boldsymbol{R}^i[y]}{\boldsymbol{R}^i[x^{(k)}]}\right)^{-\frac{n}{2}} \exp\left[-\frac{n}{2}\mathrm{tr}\big[\big(\boldsymbol{R}^i[y]^{-1} - \boldsymbol{R}^i[x^{(k)}]^{-1}\big)\boldsymbol{B}\big]\right] \frac{T(y|\alpha_i, \beta_i)}{T(x^{(k)}|\alpha_i, \beta_i)}\right\}.$$

(4) Take $x^{(k+1)} = \begin{cases} y, & \text{with probability } \delta, \\ x^{(k)}, & \text{else.} \end{cases}$

(5) If $k = I_{\mathrm{MH}} - 1$ then stop and set $z_i = x^{(k+1)}$, else set $k = k+1$ and go to (2).

(III) Set $(r_1^{(t+1)}, \ldots, r_6^{(t+1)}) = (z_1, \ldots, z_6)$.
If $t = I_{\mathrm{Gibbs}} - 1$ then stop, else set $t = t+1$ and go to (II).

□

Finally, we want to remark that a simulation for the weakly informative uniform prior (see Example 3.1) can be done in a similar way as in the examples above and it is not necessary to provide details about the algorithm.

Correlations		Moments		Triangular priors		Beta priors	
i		μ_i	σ_i	α_i	β_i	α_i	β_i
1	MR-CR	0.58	0.067	0.416	0.744	30.894	22.372
2	MR-OR	0.35	0.06	0.203	0.497	21.768	40.426
3	MR-BR	0.65	0.065	0.491	0.809	34.350	18.496
4	CR-OR	0.25	0.06	0.103	0.379	12.771	38.313
5	CR-BR	0.6	0.067	0.436	0.764	31.478	20.986
6	OR-BR	0.68	0.067	0.516	0.844	32.282	15.192

Table 4.2: Mean values and standard deviations for the six pair correlations $r_i, i = 1, \ldots, 6$, as it could be obtained by expert judgement. The associated hyperparameters (α_i, β_i) for the beta model of Example 3.2 and the triangular model of Example 3.3 are derived from (3.5) and (3.7), respectively.

4 A simulation study of aggregated EC

We now illustrate our new approach by means of a fictitious numerical example. We assume that the empirical correlation matrix $\boldsymbol{B}$ of the Gaussian copula as defined in (2.4) is given by

$$\boldsymbol{B} = \left(\begin{array}{c|cccc} & \text{MR} & \text{CR} & \text{OR} & \text{BR} \\ \hline \text{MR} & 1 & 0.66 & 0.30 & 0.58 \\ \text{CR} & 0.66 & 1 & 0.30 & 0.67 \\ \text{OR} & 0.30 & 0.30 & 1 & 0.60 \\ \text{BR} & 0.58 & 0.67 & 0.60 & 1 \end{array} \right), \tag{4.1}$$

which are actually the benchmark inter-risk correlations reported in the IFRI/CRO survey [15], Figure 10. This matrix was also used in the simulation study in Böcker [5]. Since this matrix is not derived from actual risk proxy data, we have to assume a fictitious value for the time series length n in the likelihood function (2.3). We set $n = 12$ and 72 to analyse the impact of different data length on the final results.

In addition to the empirical information above we assume subjective prior knowledge in order to completely specify the posterior distribution (3.8). Let us suppose that expert elicitation as explained in the next Section has been performed to estimate the mean values μ_i and standard deviations σ_i of all pair correlations $r_i, i = 1, \ldots, 6$, which by relationships (3.5) and (3.7) can be used to compute the hyperparameters α_i and β_i. The assumed outcome of the expert judgement is shown in Table 4.2 for all six pair correlations. Moreover, Figures 4-7 in the Appendix depict the prior densities parameterised according to Table 4.2 together with the empirical estimates (4.1).

For the simulation of the posterior distribution of the inter-risk-correlation matrix we

n=12		MR-CR	MR-OR	MR-BR	CR-OR	CR-BR	OR-BR
Uniform	Posterior Mean	0.502	0.264	0.435	0.257	0.527	0.484
	Posterior Std.	0.168	0.151	0.169	0.150	0.163	0.163
Beta	Posterior Mean	0.590	0.354	0.624	0.257	0.599	0.645
	Posterior Std.	0.061	0.056	0.055	0.057	0.058	0.055
Triangular	Posterior Mean	0.589	0.354	0.623	0.249	0.597	0.643
	Posterior Std.	0.063	0.057	0.054	0.054	0.059	0.053
n=72							
Uniform	Posterior Mean	0.626	0.267	0.543	0.266	0.640	0.570
	Posterior Std.	0.063	0.095	0.072	0.093	0.059	0.068
Beta	Posterior Mean	0.616	0.339	0.598	0.269	0.624	0.620
	Posterior Std.	0.046	0.047	0.042	0.050	0.043	0.041
Triangular	Posterior Mean	0.617	0.337	0.596	0.264	0.623	0.617
	Posterior Std.	0.047	0.048	0.041	0.047	0.044	0.039

Table 4.3: Mean values and standard deviations of the simulated marginal posterior distributions for the six pair correlations $r_i, i = 1, \ldots, 6$, assuming three different prior distributions and two different lengths for the empirical time series. One can see that, specifically for smaller n, the standard deviation of the correlations is significantly reduced when prior knowledge is incorporated.

ran the MCMC algorithms described in the previous section for the uniform,the beta and the triangular model. The starting value $\boldsymbol{R}^{(0)}$ was chosen as the empirical matrix (4.1). Furthermore, for all models we set $I_{\mathrm{MH}} = 1$ and $I_{\mathrm{Gibbs}} = 10^7$. From this chain we took only every one hundredth value in order to reduce autocorrelation; moreover, we dropped the first 10,000 as a burn-in sample. Hence, in both models we finally came up with 90,000 inter-risk-correlation matrices sampled from the posterior distribution.

Table 4.3 shows the mean values and the standard deviations of the marginal posterior distributions. The posterior means are different from the pure empirical estimates given in (4.1) because of the additional consideration of expert prior knowledge. Moreover, we see that the posterior statistics depend only very weakly on whether beta or triangular priors are chosen. Figures 4-7 graphically compare for $n = 12$ and 72 the marginal posterior distributions with the pairwise prior distributions for the beta and triangular model.

To aggregate EC we now use the Gaussian copula model, however, instead of applying a fixed correlation matrix, we use different correlation matrices randomly selected from the simulated Markov chain (after discarding the burn-in sample and adjustment for autocorrelation). In doing so, we are able to take the uncertainty of the inter-risk-correlation matrix correctly into account. Results for the beta and uniform prior models are given

Aggregated Economic Capital at 99.95 % CL		
Sum	Uniform prior	Beta prior
100	76.68 [67.86, 83.30]	78.29 [75.94, 80.47]

Table 4.4: Aggregated EC at confidence level of 99.95 % together with the 95 % credible intervals for a Bayesian Gaussian copula model using uniform and beta priors and time series of length $n = 12$. The credible intervals correspond to relative uncertainties of about 20 % and 6 % for the uniform and beta model, respectively. The portfolio consisting of market, credit, operational, and business risk is specified in Table 2.1.

in Table 4.4. One observes that beta priors lead to slightly higher results than uniform priors, which is in accordance with the posterior means shown in Table 4.3. The posterior distribution of the inter-risk-correlation matrix implies also a posterior distribution for the aggregated EC which can be used to analyse the uncertainty of a bank's aggregated EC figure and thus also the diversification benefit due to risk-type dependence. Particularly useful are graphical methods that depict the posterior density of the aggregated EC at confidence level κ, denoted by $p_{\mathrm{EC}(\kappa)}(\cdot)$, as shown in Figure 1 or "uncertainty" plots like in Figure 2. The latter illustrates the density $p_{\mathrm{EC}(\kappa)}(\cdot)$ as a function of the confidence level κ by means of a gray-level intensity plot. Other useful measures of uncertainty are credible intervals, which are direct probability statements about model parameters or functions of it given the observed data (see e.g. Berger [4]). We calculated the 95 % percent credible interval for the EC distributions given in Figure 1. In case of the beta priors we obtain $[36.63, 39.24]$ and for the triangular shaped priors a calculation yields $[36.59, 39.27]$. Obviously, in our numerical example, one can see from Table 4.4 and Figure 1 that the impact of differently shaped priors (with equal means and standard deviations, however) on the aggregated EC can be neglected.

5 Expert judgement and subjective prior assessment

This Section is devoted to the selection of appropriate prior distributions for the correlation matrix $\boldsymbol{R}$ of the Gaussian copula. In typical risk-aggregation problems data are scarce and therefore it is worthwhile to study how available subjective information or prior beliefs about inter-risk correlations can be accounted for in a formal and sound way.

The necessity for expert judgement Clearly, if almost perfect empirical data were available (e.g. complete, reliable, and representative risk-proxy time series for each risk

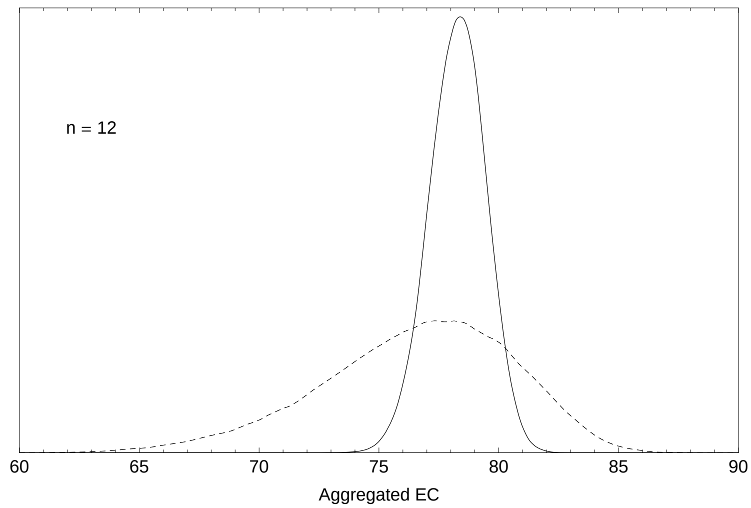

Figure 1: Posterior distribution $p_{\mathrm{EC}(\kappa)}$ of the aggregated EC at confidence level of $\kappa = 99.95$ % for the beta model (solid line) and the uniform model (dashed line) in case of $n = 12$ smoothed with Epanechnikov kernel density estimation. The posterior means and 95 % percent credible intervals are given in Table 4.4.

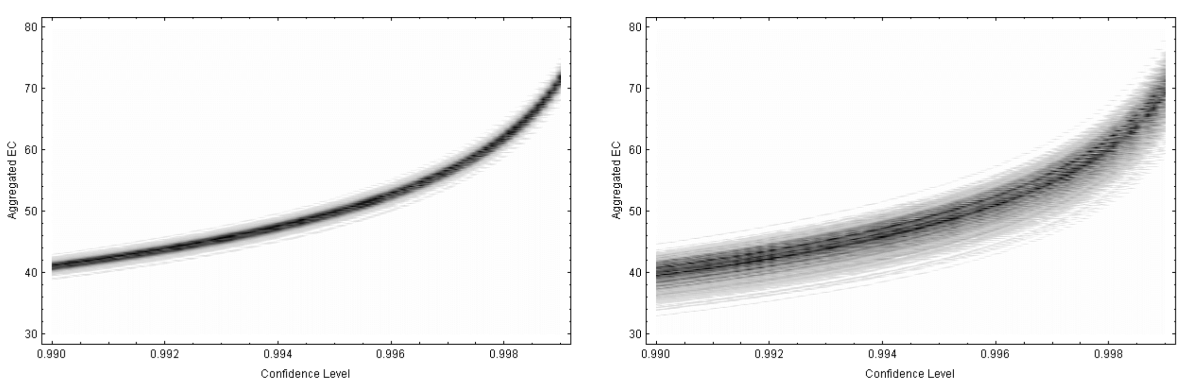

Figure 2: Uncertainty plots for the aggregated EC if the priors for the pair correlations are described by triangular distributions (left figure) and $[0, 1]$-uniform distributions(right figure). In both cases it is $n = 12$. Lighter (darker) shaded regions of the plot indicate a lower (higher) posterior density of the aggregated EC.

type) then it would be acceptable to rely only on statistical correlation estimates to approximate inter-risk dependence. Unfortunately, in practice it is often extremely difficult to identify and gather high-quality risk proxies for each risk type, making the estimation of inter-risk correlations a tricky exercise.

There are several reasons for these difficulties. First, there is the question regarding internal versus external data. The general opinion is that risk proxies should be derived from bank-internal data because it is more related to the company's specific business strategy. However, internal time series may be hard to come by or difficult to re-build after a merger or a re-organisation of its business structure, resulting in quite short risk-proxy time series, which reduce the statistical significance of the correlation estimates. A consequence thereof is that bank-internal proxies are often amended by external data, at least for some risk types. Another problem is to find a common frequency that provides a natural scale for all risk types. Usually, risk proxies for different risk types are measured at unequal time intervals and therefore further assumptions and approximations have to be accepted to make risk proxies comparable. For example, market risk data are available at a daily basis whereas proxies for business risk are often based on accounting information and therefore exist only at quarterly or even yearly level, creating a bottleneck for the statistical correlation estimation.

These drawbacks of a purely statistical analysis based on risk proxies show that it is necessary to include a new component of information, namely expert judgement. The usage of some kind of judgemental approaches when assessing inter-risk correlations is very popular in the banking industry, however, most of the employed approaches lack a sound scientific basis. For instance, the widely used method of "ex-post adjustments" of the statistical correlation estimates cannot be properly formalised, thereby nourishing fears and doubts concerning the final figures.

In contrast to this, the Bayesian approach for risk aggregation we are proposing here allows treating empirical data on the one hand and expert knowledge on the other as two distinct sources of information, which are eventually amalgamated by means of Bayes theorem. In this way it is possible to exploit both types of information without negative feedback effects, namely experts that are biased ("anchored") by the data and, vice versa, data that is manipulated by the expert.

The Bayesian choice for risk aggregation means that the experts' beliefs about the association between different risk types have to be encoded in the pairwise correlation priors $\pi_i(\cdot)$ for $i = 1, \ldots, 6$, introduced earlier in this Chapter. This process is referred to as *elicitation* of the prior distributions. Elicitation is a difficult task and a number of different competencies are required to perform it correctly, not only in statistics but also in the field of psychology. A readable textbook on this intriguing subject is O'Hagan et al. [19]. Without going into detail, we want to mention that people are typically employing only a few strategies or heuristics to quantify uncertainty or to make decisions under uncertainty, see e.g. Tversky & Kahneman [25]. Consequently, the way questions are asked and how answers are interpreted by the facilitator is crucial, in particular, it is advisable to collect expert judgements stemming from different elicitation approaches to be able to double check their internal consistency. Therefore, our suggestion here is to elicit pairwise correlations by asking questions about the following three variables, which then can be used to compute the related copula parameters.

(1) Kendall's tau rank correlations between two distinct risk types,

(2) conditional loss probabilities between two distinct risk types,

(3) joint loss probabilities between two distinct risk types.

Since we elicit single pair correlations we do not account for positive semidefiniteness of the entire correlation matrix, which therefore will be considered during the MCMC simulation of the posterior distribution.

Correlation elicitation using Kendall's tau Asking experts to quantify an inter-risk correlation directly is not a trivial task. Direct assessment of a dependence measure requires deep knowledge of the relative behaviour of each pair of variables or, in our case, of two risk types. Moreover, direct correlation estimation should be supported by a thorough explanation about which kind of correlation coefficient one is actually interested in.

The most common correlation measure used in practice is Pearson's linear correlation coefficient. However, it is well-known that this measure is not consistent with Gaussian copula risk aggregation unless the joint risk-type distribution is multivariate elliptical, see e.g. Embrechts, McNeil & Straumann [11]. Furthermore, it is by far not clear whether Pearson's linear correlation is really the kind of association measure people are thinking of when being asked about "correlations". With this respect, our main concern is that the liner correlation $\mathrm{corr}(X, Y)$ between two risk types X and Y depends not only on their dependence structure but also on the specific form of the marginal distribution functions F_X and F_Y. Therefore, experts will only interpret and estimate a linear correlation correctly if they also account for the specific marginal distribution functions assumed for X and Y. Another problem is the counter-intuitive fact that for given risk-type marginals F_X and F_Y the attainable correlations lie, in general, in a subinterval of $[-1, 1]$. All these problems have to be clarified before beginning the elicitation experiment since naive, ambiguous questions about some kind of risk-type "correlation" may lead the expert to give an answer biased by her cognitive notion of correlation that probably significantly differs from the Gaussian copula parameter we are actually interested in.

A possible loophole is an alternative concept of association, namely that of Kendall's tau rank correlation τ, which is particularly useful when—as in our case—a multivariate elliptical problem is assumed, cf. for instance Embrechts, Lindskog & McNeil [10], and with an application to risk-type aggregation and elicitation Böcker [5]. The benefit for expert elicitation is due to the relationship between Kendall's tau τ_i and the associated Gaussian copula parameter r_i, which holds true for essentially all elliptical distributions, namely

$$r_i = \sin(\pi\,\tau_i/2)\,, \qquad i = 1, \ldots, 6\,. \tag{5.1}$$

Recall that $\tau_i \in [-1, 1]$, with $\tau_i = 1, (\tau_i = -1)$ for complete positive (negative) dependence, and $\tau_i = 0$ for independent risk types. Our suggestion is now to ask experts about their correlation estimates within an interval $[-1, 1]$ where $\tau_i = \{-1, 0, 1\}$ can be used as anchors helping experts to calibrate their answers. Finally, the elicited value for τ_i can be transformed to the copula parameter r_i by means of (5.1).

Correlation elicitation using conditional and joint probabilities The correlation assessment described above requires the expert to think about two random variables simultaneously and is an example for a bivariate elicitation task. A natural way to lower the complexity of the elicitation procedure is to reduce it to a univariate problem. Specifically, it has been argued in Gokhale & Press [13] or O'Hagan et al. [19] that one-dimensional problems are more feasible for expert judgement as experts elicit univariate variables with a higher degree of accuracy.

Correlation elicitation by means of indirect questions about conditional and joint probabilities was already described in Böcker [5] so that here we only briefly summarise the main results. First note that if d risk types are jointly distributed with a Gaussian copula with correlation matrix $(R_{ij})_{ij}, i, j = 1, \ldots, d$, any two risk types l and m $(m \neq l)$ are coupled by a Gaussian copula with correlation parameter R_{lm}. Hence it is sufficient to consider the bivariate estimation problem.

For two risk-type variables X and Y (and the definition that losses are positive) we can express their joint survival probability as

$$\begin{aligned} P(X > x, Y > y) &= 1 - P(X \leq x) - P(Y \leq y) + P(X \leq x, Y \leq y) \\ &= 1 - F_X(x) - F_Y(y) + C_R(F_X(x), F_Y(y))\,, \quad x, y > 0\,, \end{aligned} \tag{5.2}$$

where F_X and F_Y are the marginal distribution functions of X and Y, respectively, and $C_R(\cdot, \cdot)$ is the bivariate Gaussian copula with unknown parameter R. Similarly, one may

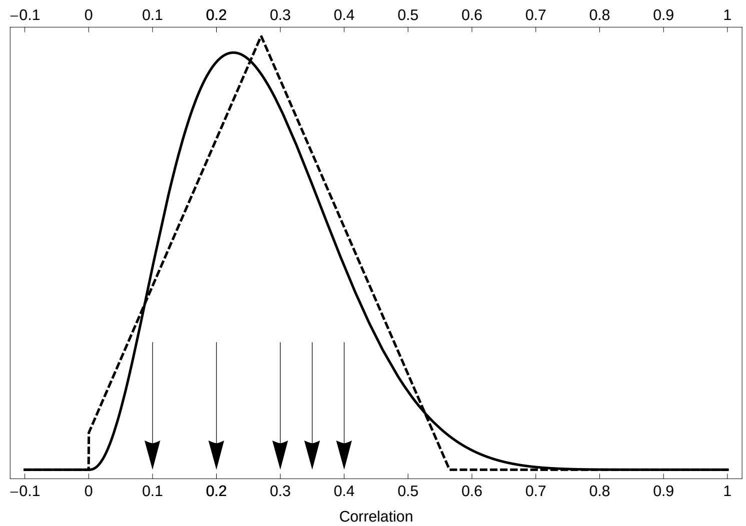

Figure 3: Beta prior and triangular prior obtained by moment matching. It is assumed that the experts' estimates of some pair correlation are $r_i^* = \{0.1, 0.2, 0.3, 0.35, 0.4\}$, yielding $\overline{r_i^*} = 0.27$ and $\mathrm{var}(r_i^*) = 0.0145$.

consider conditional loss probabilities of the form

$$P(X > x|Y > y) = \frac{P(Y > y, X > x)}{P(Y > y)}, \quad x, y > 0\,, \tag{5.3}$$

which depend on the copula correlation R via (5.2). Now, our strategy is to elicit such joint and conditional probabilities for different threshold values of x and y. Since we assume that the marginal distribution functions F_X and F_Y are known and already completely parameterised, one can use relationships (5.2) and (5.3) to determine the copula correlation R, usually by numerical of graphical methods.

The result of the entire expert elicitation program is that for each pairwise correlation $r_i, i = 1, \ldots, 6$, one obtains a sample $r_i^* := \{r_{i1}^*, \ldots, r_{iN_i}^*\}$ of N_i expert estimates. The elicited values r_i^* may arise from one expert who answered to all of the three approaches suggested above, or may reflect different opinions about risk-type dependence stemming from several different experts. Now, the experts' point estimates $r_i^*, i = 1, \ldots, 6$, can be used to determine the hyperparameters of the pairwise correlation priors $\pi_i(\cdot), i = 1, \ldots, 6$. For the beta and triangular models these are only two parameters α_i and β_i for each pair correlation. Therefore, a viable approach is moment matching because, as shown in Examples 3.2 and 3.3, we only need to decide about the prior means and variances to fully determine all $\pi_i(\cdot), i = 1, \ldots, 6$. This approach is illustrated in Figure 3.

6 Acknowledgement

Thanks are due to Alexander Thomas for essential support in performing the simulations used in this approach.

7 Figures

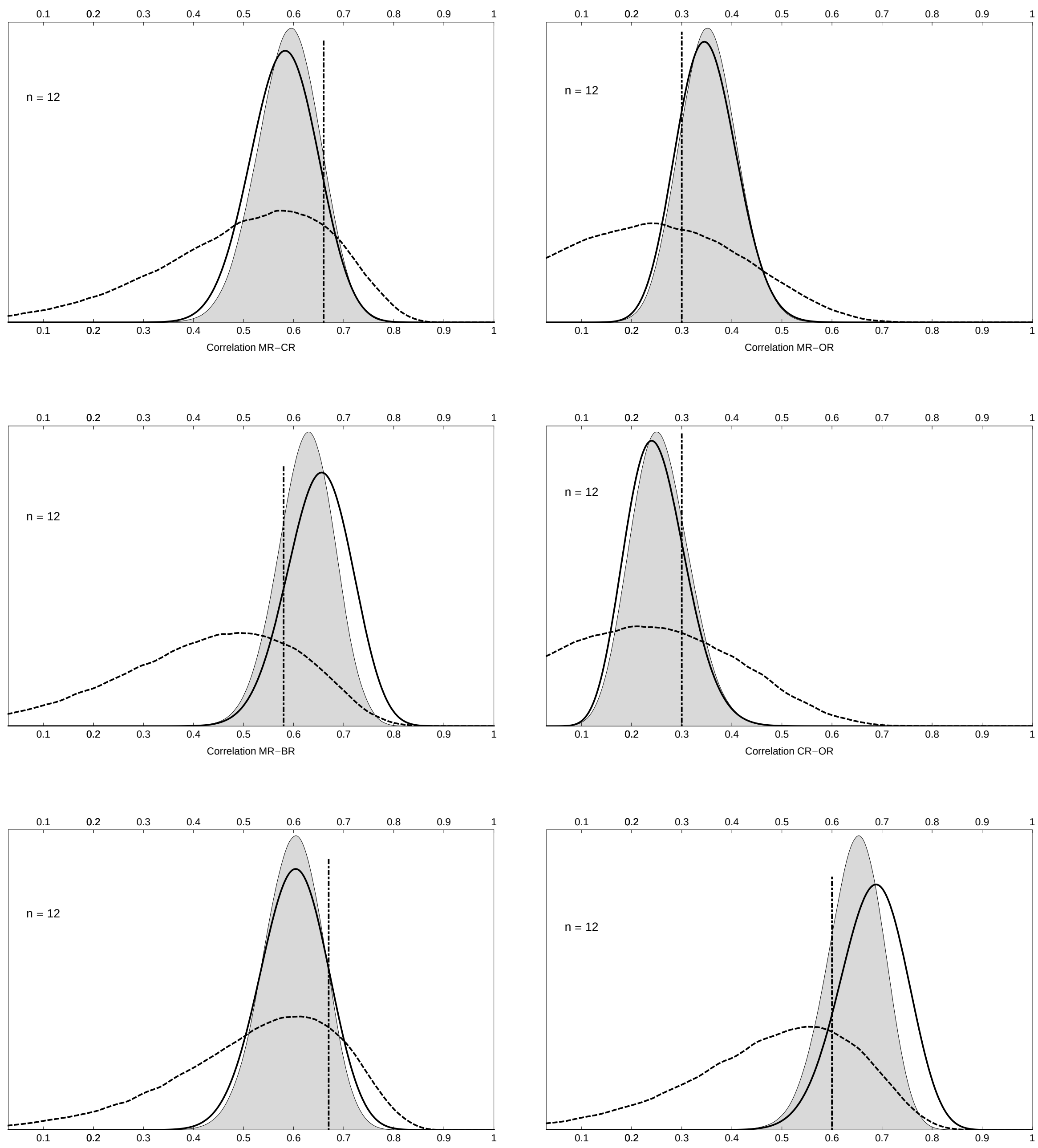

Figure 4: Plots of the simulated marginal posterior distributions (kernel density estimation with an Epanechnikov kernel) and the beta distributed priors for the six pair correlations $r_i, i = 1, \ldots, 6$ for $n = 12$. The grey filled curves are the marginal posteriors obtained from the beta priors (solid curves), the dashed curves result from using $[0, 1]$-uniformly distributed priors. The dotted-dashed lines indicate the empirical correlations in (4.1).

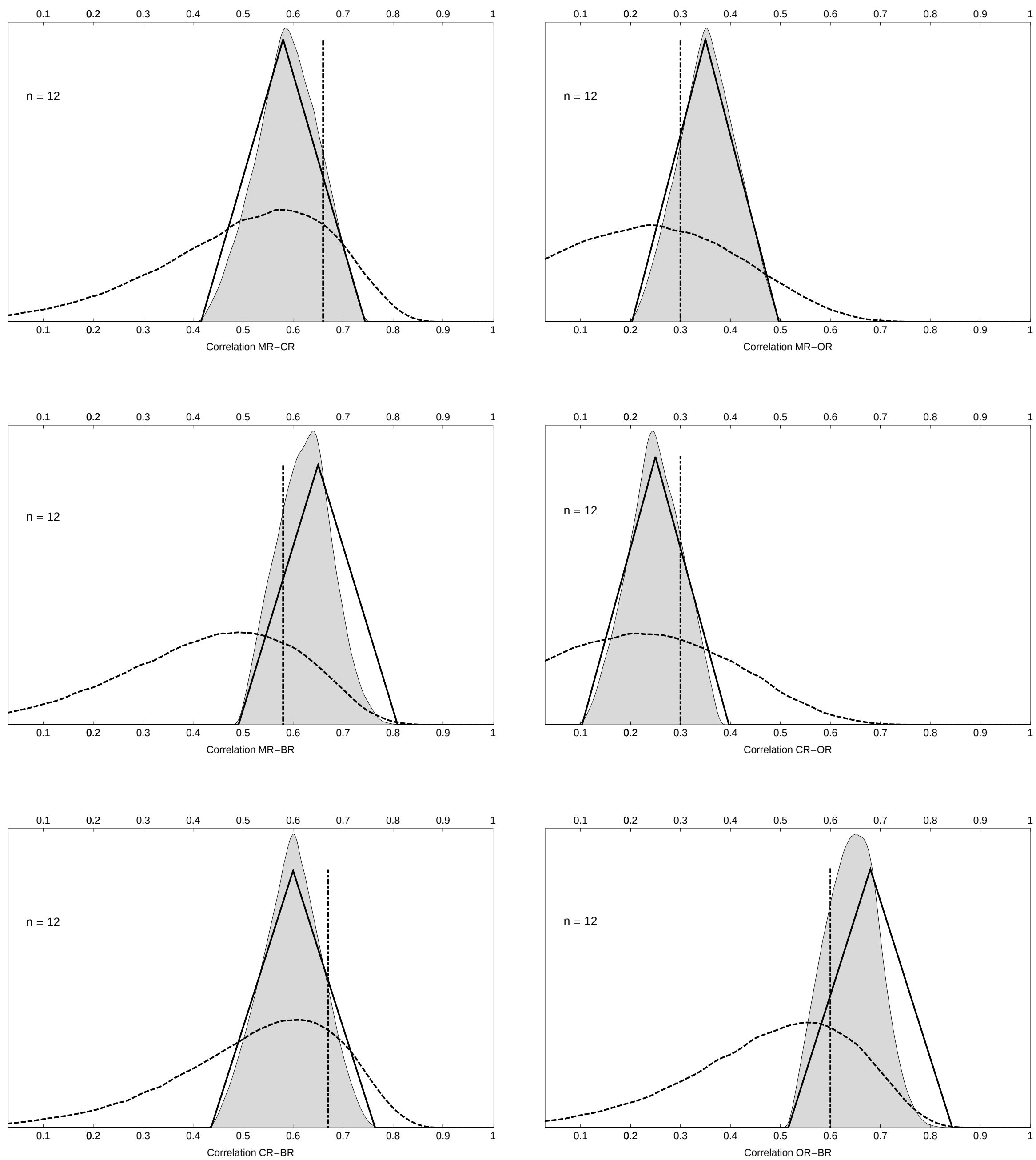

Figure 5: Plots of the simulated marginal posterior distributions (kernel density estimation with an Epanechnikov kernel) and the triangular distributed priors for the six pair correlations $r_i, i = 1, \ldots, 6$ for $n = 12$. The grey filled curves are the marginal posteriors obtained from the triangular priors (solid curves), the dashed curves result from using $[0, 1]$-uniformly distributed priors. The dotted-dashed lines indicate the empirical correlations in (4.1).

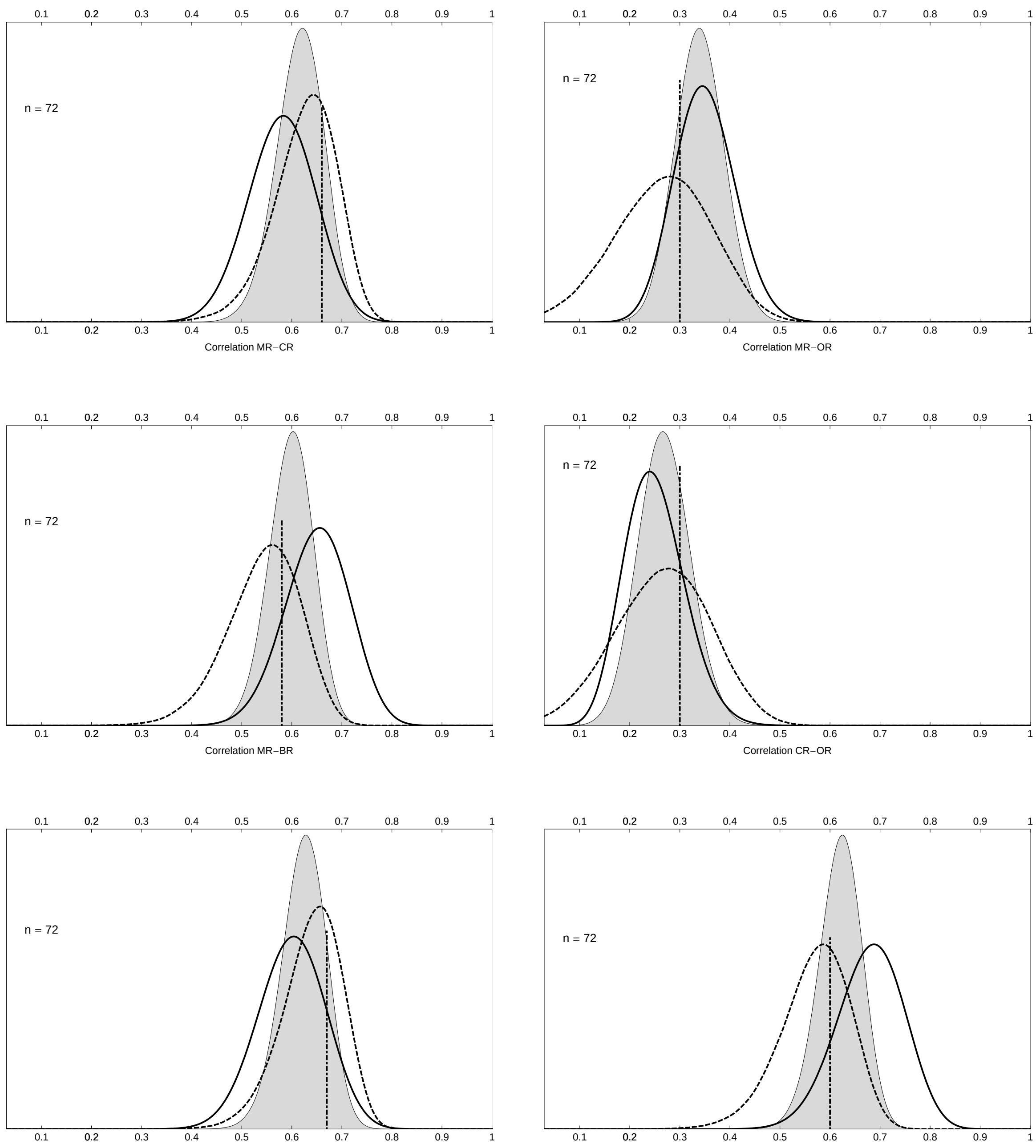

Figure 6: Plots of the simulated marginal posterior distributions (kernel density estimation with an Epanechnikov kernel) and the beta distributed priors for the six pair correlations $r_i, i = 1, \ldots, 6$ for $n = 72$. The grey filled curves are the marginal posteriors obtained from the beta priors (solid curves), the dashed curves result from using $[0, 1]$-uniformly distributed priors. The dotted-dashed lines indicate the empirical correlations in (4.1).

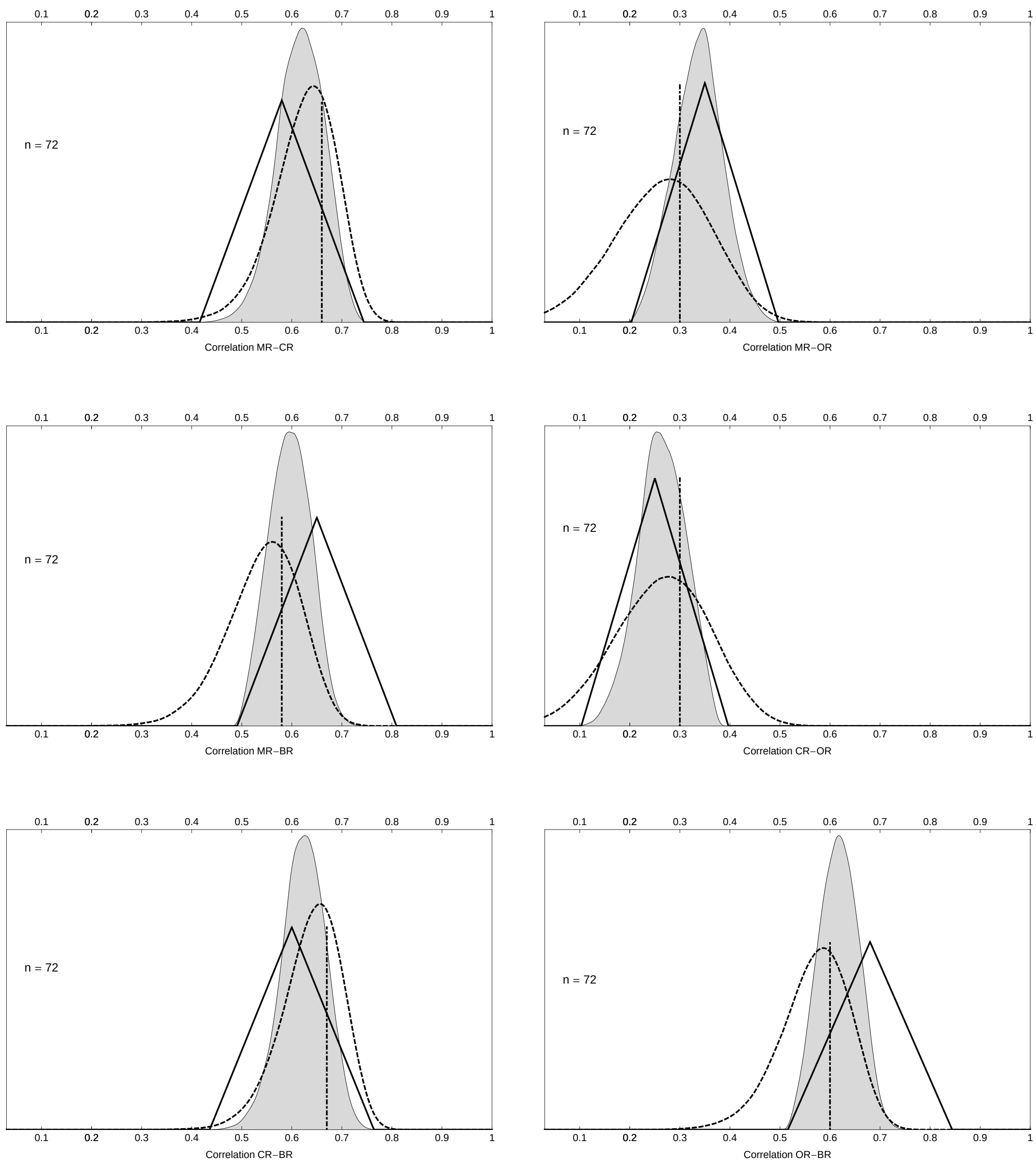

Figure 7: Plots of the simulated marginal posterior distributions (kernel density estimation with an Epanechnikov kernel) and the triangular distributed priors for the six pair correlations $r_i, i = 1, \ldots, 6$ for $n = 72$. The grey filled curves are the marginal posteriors obtained from the triangular priors (solid curves), the dashed curves result from using $[0, 1]$-uniformly distributed priors. The dotted-dashed lines indicate the empirical correlations in (4.1).

References

[1] J. Barnard, R. McCulloch, X.-L. Meng, Modeling covariance matrices in terms of standard deviations and correlations, with application to shrinkage. Statistica Sinica 10, 1281-1311, (2000)

[2] Basel Committee on Banking Supervision (BCBS), Findings on the interaction of market and credit risk, Working Paper 16, BIS. (2009)

[3] Basel Committee on Banking Supervision (BCBS), Range of practices and issues in economic capital frameworks, BIS. (2009)

[4] J. O. Berger, *Statistical Decision Theory and Bayesian Analysis*, 2nd ed., Springer-Verlag, New York, (1985)

[5] K. Böcker, Aggregation by Risk Type and Inter Risk Correlation, In: *Pillar II in the New Basel Accord – The Challenge of Economic Capital*, Ed. A. Resti, Risk Books, London, 313-344, (2008)

[6] M. Brockmann, M. Kalkbrener, *On the Aggregation of Risk*, Working Paper, Deutsche Bank, (2008)

[7] U. Cherubini, E. Luciano, W. Vecchiato, *Copula methods in finance*, John Wiley & Sons, Chichester, West Sussex, (2004)

[8] Committee of European Banking Supervisors (CEBS): Consultation paper on technical aspects of diversification under Pillar 2 (CP20), Consultation Paper, (2008)

[9] X. K. Dimakos, K. Aas, Integrated Risk Modelling, Statistical Modelling **4**, 265-277, (2004)

[10] P. Embrechts, F. Lindskog, A. McNeil, *Modelling Dependence with Copulas and Applications to Risk Management*, In: Handbook of Heavy Tailed Distributions in Finance, Ed. S. Rachev, Elsevier, 329-384, (2001)

[11] P. Embrechts, A. McNeil, D. Straumann, Correlation and dependence in risk management: properties and pitfalls, In: *Risk Management: Value at Risk and Beyond*, Ed. M.A.H. Dempster, Cambridge University Press, Cambridge, 176-223, (2002)

[12] W. R. Gilks, S. Richardson, D. J. Spiegelhalter (Eds.), *Markov Chain Monte Carlo in Practice*, Chapman & Hall, London (1996)

[13] D. V. Gokhale, S. James Press, Assessment of a Prior Distribution for the Correlation Coefficient in a Bivariate Normal Distribution, Journal of the Royal Statistical Society A, 145(**2**), 237-249, (1982)

[14] P. Hartmann, M. Pritsker, T. Schuermann (Eds.), Interaction of market and credit risk, Special Issue of the Journal of Banking and Finance, forthcoming (2009)

[15] The Institute of the Chief Risk Officers/Chief Risk Officer Forum, *Insights from the joint IFRI/CRO Forum survey on Economic Capital practice and applications*, (2007)

[16] G. Kretzschmar, A.J. McNeil, A. Kirchner, Integrated models of capital adequacy Why banks are undercapitalised, Technical Report, (2009)
Available at http://www.ma.hw.ac.uk/ mcneil/ftp/WhyBanksUndercapitalised.pdf

[17] P. Müller, A generic approach to posterior integration and Gibbs sampling, Technical Report, Purdue Univ., West Lafayette, Indiana, (1991)

[18] P. Müller, Alternatives to the Gibbs sampling scheme, Technicl Report, Institute of Statistics and Decision Sciences, Duke Univ., (1993)

[19] A. O'Hagan et al., Uncertain Judgements: Eliciting Experts' Probabilities, John Wiley & Sons, Chichester, West Sussex, (2006)

[20] S. J. Press, *Applied multivariate analysis: using Bayesian and frequentist methods of inference*, 2nd ed., Dover Publications, Mineola, New York, (2005)

[21] C. P. Robert, G. Casella, *Monte Carlo Statistical Methods*, 2nd ed., Springer, New York, (2004)

[22] G. Roberts, Markov chain concepts related to sampling algorithms, In: *Markov Chain Monte Carlo in Practice*, Ed.: W. R. Gilks, S. Richardson, D. J. Spiegelhalter, Chapman & Hall, London, 45-57, (1996)

[23] J. V. Rosenberg, T. Schuermann, A General Approach to Integrated Risk Management with Skewed, Fat-Tailed Risk, Working Paper, Federal Reserve Bank of New York Staff Reports, no. 185, (May 2004)

[24] F. Saita, Principles of Risk Aggregation, In: *Pillar II in the New Basel Accord – The Challenge of Economic Capital*, Ed. A. Resti, Risk Books, London, 297-311, (2008)

[25] A. Tversky, D. Kahneman, Judgment under Uncertainty: Heuristics and Biases, Science, (**185**), 1124-1131, (1974)

[26] L. Ward, D. Lee, Practical application of risk-adjusted return on capital framework, Dynamic Financial Analysis Discussion Paper, CAS Forum, (2002)