

Article

# Markov Chain Monte Carlo Methods for Estimating Systemic Risk Allocations

Takaaki Koike \*  and Marius Hofert 

Department of Statistics and Actuarial Science, University of Waterloo, 200 University Avenue West, Waterloo, ON N2L 3G1, Canada; marius.hofert@uwaterloo.ca

\* Correspondence: tkoike@uwaterloo.ca

Received: 24 September 2019; Accepted: 7 January 2020; Published: 15 January 2020



**Abstract:** In this paper, we propose a novel framework for estimating systemic risk measures and risk allocations based on Markov Chain Monte Carlo (MCMC) methods. We consider a class of allocations whose  $j$ th component can be written as some risk measure of the  $j$ th conditional marginal loss distribution given the so-called crisis event. By considering a crisis event as an intersection of linear constraints, this class of allocations covers, for example, conditional Value-at-Risk (CoVaR), conditional expected shortfall (CoES), VaR contributions, and range VaR (RVaR) contributions as special cases. For this class of allocations, analytical calculations are rarely available, and numerical computations based on Monte Carlo (MC) methods often provide inefficient estimates due to the rare-event character of the crisis events. We propose an MCMC estimator constructed from a sample path of a Markov chain whose stationary distribution is the conditional distribution given the crisis event. Efficient constructions of Markov chains, such as the Hamiltonian Monte Carlo and Gibbs sampler, are suggested and studied depending on the crisis event and the underlying loss distribution. The efficiency of the MCMC estimators is demonstrated in a series of numerical experiments.

**Keywords:** systemic risk measures; conditional Value-at-Risk (CoVaR); capital allocation; copula models; quantitative risk management

## 1. Introduction

In portfolio risk management, *risk allocation* is an essential step to quantifying the risk of each unit of a portfolio by decomposing the total risk of the whole portfolio. One of the most prevalent rules to determine risk allocations is the *Euler principle*, proposed by [Tasche \(1995\)](#) and justified from various viewpoints, such as the RORAC compatibility ([Tasche \(1995\)](#) and [Tasche \(2008\)](#)) and cooperative game theory ([Denault \(2001\)](#)). For the popular risk measures, such as VaR, RVaR, and ES, Euler allocations take the form of conditional expectations of the underlying loss random vector given a certain rare event on the total loss of the portfolio; see [Tasche \(2001\)](#) for derivations. We call this rare event the *crisis event*. The decomposition of risks is also required in the context of systemic risk measurement. *Systemic risk* is the risk of financial distress of an entire economy as a result of the failure of individual components of the financial system. To quantify such risks, various *systemic risk measures* have been proposed in the literature, such as *conditional VaR (CoVaR)* ([Adrian and Brunnermeier \(2016\)](#)), *conditional expected shortfall (CoES)* ([Mainik and Schaanning \(2014\)](#)), and *marginal expected shortfall (MES)* ([Acharya et al. \(2017\)](#)). These three measures quantify the risk of individuals by taking the VaR, ES, and expectation of the individual loss, respectively, under some stressed scenario—that is, given the crisis event. [Chen et al. \(2013\)](#), [Hoffmann et al. \(2016\)](#), and [Kromer et al. \(2016\)](#) proposed an axiomatic characterization of systemic risk measures, where the risk of the aggregated loss in a financial system is first measured and then decomposed into the individual economic entities. Due to the similarity of risk allocations with the derivation of systemic risk measures, we refer to both of them as *systemic risk*

*allocations*. In fact, MES coincides with the Euler allocation of ES, and other Euler allocations can be regarded as special cases of systemic risk measures considered in [Gourieroux and Monfort \(2013\)](#).

Calculating systemic risk allocations given an unconditional joint loss distribution is generally challenging, since analytical calculations often require knowledge of the joint distribution of the marginal and aggregated loss. Furthermore, MC estimation suffers from the rare-event character of the crisis event. For computing CoVaR, CoES, and MES, [Mainik and Schaanning \(2014\)](#), [Bernardi et al. \(2017\)](#), and [Jaworski \(2017\)](#) derived formulas based on the copula of the marginal and aggregated loss; [Asimit and Li \(2018\)](#) derived asymptotic formulas based on the extreme value theory; and [Girardi and Ergün \(2013\)](#) estimated CoVaR under a multivariate GARCH model. [Vernic \(2006\)](#), [Chiragiev and Landsman \(2007\)](#), [Dhaene et al. \(2008\)](#), and [Furman and Landsman \(2008\)](#) calculated Euler allocations for specific joint distributions. [Asimit et al. \(2011\)](#) derived asymptotic formulas for risk allocations. [Furman and Zitikis \(2009\)](#) and [Furman et al. \(2018\)](#) calculated weighted allocations, which include Euler allocations as special cases, under a Stein-type assumption. Concerning the numerical computation of Euler allocations, [Glasserman \(2005\)](#), [Glasserman and Li \(2005\)](#), and [Kalkbrener et al. \(2004\)](#) considered importance sampling methods, and [Siller \(2013\)](#) proposed the Fourier transform Monte Carlo method, all specifically for credit portfolios. For general copula-based dependence models, analytical calculations of systemic risk allocations are rarely available, and an estimation method is, to the best of our knowledge, only addressed in [Targino et al. \(2015\)](#), where sequential Monte Carlo (SMC) samplers are applied.

We address the problem of estimating systemic risk allocations under general copula-based dependent risks in the case where the copula between the marginal and aggregated losses are not necessarily available. We consider a general class of systemic risk allocations in the form of risk measures of a conditional loss distribution given a crisis event, which includes CoVaR, CoES, MES, and Euler allocations as special cases. In our proposed method, the conditional loss distribution, called the *target distribution*  $\pi$ , is simulated by a Markov chain whose stationary distribution is the desired distribution  $\pi$  by sequentially updating the sample path based on the available information from  $\pi$ . While this MCMC method resembles the SMC in [Targino et al. \(2015\)](#), the latter requires a more complicated implementation involving the choice of forward and backward kernels, resampling and move steps, and even MCMC in the move steps. Our suggested approach directly constructs a single sophisticated Markov chain depending on the target distribution of interest. Applications of MCMC to estimating risk allocations have been studied in [Koike and Minami \(2019\)](#), specifically for VaR contributions. Our paper explores and demonstrates the applicability of MCMC methods to a more general class of systemic risk allocations.

Almost all MCMC methods used in practice are of the *Metropolis–Hastings (MH)* type ([Metropolis et al. \(1953\)](#) and [Hastings \(1970\)](#)), where the so-called *proposal distribution*  $q$  generates a candidate of the next state based on the current state. This candidate is then accepted or rejected according to the so-called *acceptance probability* to adjust the stationary distribution to be the target distribution  $\pi$ . As explained in Section 3.1 below, the resulting Markov chain has serial correlation, which adversarially affects the efficiency of the estimator. An efficient MCMC of MH type is such that the proposal distribution generates a candidate which exhibits low correlation with the current state with sufficiently large acceptance probability. The main difficulty in constructing such an efficient MCMC estimator for systemic risk allocations is that the support of the target distribution  $\pi$  is subject to constraints determined by the crisis event. For such target distributions, simple MCMC methods, such as random walk MH, are not efficient since a candidate is immediately rejected if it violates the constraints; see Section 3.2 for details.

To tackle this problem, we consider two specific MCMC methods, *Hamiltonian Monte Carlo (HMC)* ([Duane et al. \(1987\)](#)) and the *Gibbs sampler (GS)* ([Geman and Geman \(1984\)](#) and [Gelfand and Smith \(1990\)](#)). In the HMC method, a candidate is generated according to the so-called Hamiltonian dynamics, which leads to a high acceptance probability and low correlation with the current state by accurately simulating the dynamics of sufficiently long length; see [Neal et al. \(2011\)](#) and [Betancourt \(2017\)](#) for

an introduction to HMC. Moreover, the HMC candidates always belong to the crisis event by reflecting the dynamics when the chain hits the boundary of the constraints; see Ruján (1997), Pakman and Paninski (2014), Afshar and Domke (2015), Yi and Doshi-Velez (2017), and Chevallier et al. (2018) for this reflection property of the HMC method. An alternative method to handle the constraints is the GS, in which the chain is updated in each component. Since all the components except the updated one remain fixed, a componentwise update is typically subject to weaker constraints. As long as such componentwise updates are feasible, the GS candidates belong to the crisis event, and the acceptance probability is always 1; see Geweke (1991), Gelfand et al. (1992), and Rodriguez-Yam et al. (2004) for the application of the GS to constrained target distributions, and see Gudmundsson and Hult (2014) and Targino et al. (2015) for applications to estimating risk contributions.

Our findings include efficient MCMC estimators of systemic risk allocations achieved via HMC with reflection and GSs. We assume that the unconditional joint loss density is known, possibly through its marginal densities and copula density. Depending on the supports of the marginal loss distributions and the crisis event, different MCMC methods are applicable. We find that if the marginal loss distributions are one-sided, that is, the supports are bounded from the left, then the crisis event is typically a bounded set and HMC shows good performance. On the other hand, if the marginal losses are two-sided, that is, they have both right and left tails, the crisis event is often unbounded and the GSs perform better, provided that the random number generators of the conditional copulas are available. Based on the samples generated by the MC method, we propose heuristics to determine the parameters of the HMC and GS methods, for which no manual interaction is required. Since, in the MCMC method, the conditional loss distribution of interest is directly simulated, in contrast to MC where rejection is applied based on the unconditional loss distribution, the MCMC method generally outperforms the MC method in terms of the sample size, and thus the standard error. This advantage of MCMC becomes more pronounced as the probability of the crisis event becomes smaller. We demonstrate this efficiency of the MCMC estimators of systemic risk allocations by a series of numerical experiments.

This paper is organized as follows. The general framework of the estimation problem of systemic risk allocations is introduced in Section 2. Our class of systemic risk allocations is proposed in Section 2.1, and their estimation via the MC method is presented in Section 2.2. Section 3 is devoted to MCMC methods for estimating systemic risk allocations. After a brief review of MCMC methods in Section 3.1, we formulate our problem of estimating systemic risk allocations in terms of MCMC in Section 3.2. HMC and GS for constrained target distributions are then investigated in Sections 3.3 and 3.4, respectively. In Section 4, numerical experiments are conducted, including simulation and empirical studies, and a detailed comparison of MC and our introduced MCMC methods is provided. Section 5 concludes with practical guidance and limitations of the presented MCMC methods. An R script reproducing the numerical experiments is available as Supplementary Material.

## 2. Systemic Risk Allocations and Their Estimation

In this section, we define a broad class of systemic risk allocations, including Euler allocations, CoVaR, and CoES as special cases. Then, the MC method is described to estimate systemic risk allocations.

### 2.1. A Class of Systemic Risk Allocations

Let  $(\Omega, \mathcal{F}, \mathbb{P})$  be an atomless probability space, and let  $X_1, \dots, X_d$ ,  $d \geq 2$  be random variables on this space. The random vector  $\mathbf{X} = (X_1, \dots, X_d)$  can be interpreted as losses of a portfolio of size  $d$ , or losses of  $d$  economic entities in an economy over a fixed time period. Throughout the paper, a positive value of a loss random variable represents a financial loss, and a negative loss is interpreted as a profit. Let  $F_{\mathbf{X}}$  denote the joint cumulative distribution function (cdf) of  $\mathbf{X}$  with marginal distributions  $F_1, \dots, F_d$ . Assume that  $F_{\mathbf{X}}$  admits a probability density function (pdf)  $f_{\mathbf{X}}$  with marginal densities  $f_1, \dots, f_d$ . Sklar's theorem (Nelsen (2006)) allows one to write

$$F_{\mathbf{X}}(\mathbf{x}) = C(F_1(x_1), \dots, F_d(x_d)), \quad \mathbf{x} = (x_1, \dots, x_d) \in \mathbb{R}^d, \tag{1}$$

where  $C : [0, 1]^d \rightarrow [0, 1]$  is a *copula* of  $\mathbf{X}$ . Assuming the density  $c$  of the copula  $C$  to exist,  $f_{\mathbf{X}}$  can be written as

$$f_{\mathbf{X}}(\mathbf{x}) = c(F_1(x_1), \dots, F_d(x_d))f_1(x_1) \cdots f_d(x_d), \quad \mathbf{x} \in \mathbb{R}^d.$$

An *allocation*  $A = (A_1, \dots, A_d)$  is a map from a random vector  $\mathbf{X}$  to  $(A_1(\mathbf{X}), \dots, A_d(\mathbf{X})) \in \mathbb{R}^d$ . The sum  $\sum_{j=1}^d A_j(\mathbf{X})$  can be understood as the capital required to cover the total loss of the portfolio or the economy. The  $j$ th component  $A_j(\mathbf{X}), j = 1, \dots, d$  is then the contribution of the  $j$ th loss to the total capital  $\sum_{j=1}^d A_j(\mathbf{X})$ . In this paper, we consider the following class of allocations

$$A^{q_1, \dots, q_d, \mathcal{C}} = (A_1^{q_1, \mathcal{C}}, \dots, A_d^{q_d, \mathcal{C}}), \quad A_j^{q_j, \mathcal{C}}(\mathbf{X}) = q_j(X_j | \mathbf{X} \in \mathcal{C}),$$

where  $q_j$  is a map from a random variable to  $\mathbb{R}$  called the  $j$ th *marginal risk measure* for  $j = 1, \dots, d$ , and  $\mathcal{C} \subseteq \mathbb{R}^d$  is a set called the *crisis event*. The conditioning set  $\{\mathbf{X} \in \mathcal{C}\}$  is simply written as  $\mathcal{C}$  if there is no confusion. As we now explain, this class of allocations covers well-known allocations as special cases. For a random variable  $X \sim F$ , we define the *Value-at-Risk (VaR)* of  $X$  at confidence level  $\alpha \in (0, 1]$  by

$$\text{VaR}_\alpha(X) := \inf\{x \in \mathbb{R} : F(x) \geq \alpha\}.$$

*Range Value-at-Risk (RVaR)* at confidence levels  $0 < \alpha_1 < \alpha_2 \leq 1$  is defined by

$$\text{RVaR}_{\alpha_1, \alpha_2}(X) = \frac{1}{\alpha_2 - \alpha_1} \int_{\alpha_1}^{\alpha_2} \text{VaR}_\gamma(X) d\gamma,$$

and, if it exists, *expected shortfall (ES)* at confidence level  $\alpha \in (0, 1)$  is defined by  $\text{ES}_\alpha(X) = \text{RVaR}_{\alpha, 1}(X)$ . Note that ES is also known as C(onditional)VaR, T(ail)VaR, A(verage)VaR and C(onditional)T(ail)E(xpectation). These risk measures are law-invariant in the sense that they depend only on the distribution of  $X$ . Therefore, we sometimes write  $q(F)$  instead of  $q(X)$ .

We now define various crisis events and marginal risk measures. A typical form of the crisis event is an intersection of a set of linear constraints

$$\mathcal{C} = \bigcap_{m=1}^M \left\{ \mathbf{h}_m^\top \mathbf{x} \geq v_m \right\}, \quad \mathbf{h}_m \in \mathbb{R}^d, \quad v_m \in \mathbb{R}, \quad m = 1, \dots, M, \quad M \in \mathbb{N}. \tag{2}$$

Several important special cases of the crisis event of Form (2) are provided in the following.

**Definition 1** (VaR, RVaR, and ES crisis events). For  $S = \sum_{j=1}^d X_j$ , the VaR, RVaR and ES crisis events are defined by

$$\begin{aligned} \mathcal{C}_\alpha^{\text{VaR}} &= \{ \mathbf{x} \in \mathbb{R}^d \mid \mathbf{1}_d^\top \mathbf{x} = \text{VaR}_\alpha(S) \}, \quad \alpha \in (0, 1), \\ \mathcal{C}_{\alpha_1, \alpha_2}^{\text{RVaR}} &= \{ \mathbf{x} \in \mathbb{R}^d \mid \text{VaR}_{\alpha_1}(S) \leq \mathbf{1}_d^\top \mathbf{x} \leq \text{VaR}_{\alpha_2}(S) \}, \quad 0 < \alpha_1 < \alpha_2 \leq 1, \\ \mathcal{C}_\alpha^{\text{ES}} &= \{ \mathbf{x} \in \mathbb{R}^d \mid \text{VaR}_\alpha(S) \leq \mathbf{1}_d^\top \mathbf{x} \}, \quad 0 < \alpha < 1, \quad \alpha \in (0, 1), \end{aligned}$$

respectively, where  $\mathbf{1}_d$  is the  $d$ -dimensional vector of ones.

**Definition 2** (Risk contributions and conditional risk measures). For  $j \in \{1, \dots, d\}$ , we call  $A_j^{q_j, \mathcal{C}}$  of

1. *risk contribution-type* if  $q_j = \mathbb{E}$ ;

2. CoVaR type if  $q_j = \text{VaR}_{\beta_j}$  for  $\beta_j \in (0, 1)$ ;
3. CoRVaR type if  $q_j = \text{RVaR}_{\beta_{j,1}, \beta_{j,2}}$  for  $0 < \beta_{j,1} < \beta_{j,2} \leq 1$ ; and
4. CoES-type if  $q_j = \text{ES}_{\beta_j}$  for  $\beta_j \in (0, 1)$ .

The following examples show that  $A_j^{q_j, \mathcal{C}}$  coincides with popular allocations for some specific choices of marginal risk measure and crisis event.

**Example 1** (Special cases of  $A^{q_1, \dots, q_d, \mathcal{C}}$ ).

- (1) Risk contributions. If the crisis event is chosen to be  $\mathcal{C}_\alpha^{\text{VaR}}$ ,  $\mathcal{C}_{\alpha_1, \alpha_2}^{\text{RVaR}}$  or  $\mathcal{C}_\alpha^{\text{ES}}$ , the allocations of the risk contribution type  $q_j = \mathbb{E}$  reduce to the VaR, RVaR, or ES contributions defined by

$$\begin{aligned} \text{VaR}_\alpha(\mathbf{X}, S) &= \mathbb{E}[\mathbf{X} \mid S = \text{VaR}_\alpha(S)], \\ \text{RVaR}_{\alpha_1, \alpha_2}(\mathbf{X}, S) &= \mathbb{E}[\mathbf{X} \mid \text{VaR}_{\alpha_1}(S) \leq S \leq \text{VaR}_{\alpha_2}(S)], \\ \text{ES}_\alpha(\mathbf{X}, S) &= \mathbb{E}[\mathbf{X} \mid S \geq \text{VaR}_\alpha(S)], \end{aligned}$$

respectively. These results are derived by allocating the total capital  $\text{VaR}_\alpha(S)$ ,  $\text{RVaR}_{\alpha_1, \alpha_2}(S)$  and  $\text{ES}_\alpha(S)$  according to the Euler principle; see [Tasche \(1995\)](#). The ES contribution is also called the MES and used as a systemic risk measure; see [Acharya et al. \(2017\)](#).

- (2) Conditional risk measures. CoVaR and CoES are systemic risk measures defined by

$$\begin{aligned} \text{CoVaR}_{\alpha, \beta}^-(X_j, S) &= \text{VaR}_\beta(X_j \mid S = \text{VaR}_\alpha(S)), & \text{CoVaR}_{\alpha, \beta}(X_j, S) &= \text{VaR}_\beta(X_j \mid S \geq \text{VaR}_\alpha(S)), \\ \text{CoES}_{\alpha, \beta}^-(X_j, S) &= \text{ES}_\beta(X_j \mid S = \text{VaR}_\alpha(S)), & \text{CoES}_{\alpha, \beta}(X_j, S) &= \text{ES}_\beta(X_j \mid S \geq \text{VaR}_\alpha(S)), \end{aligned}$$

for  $\alpha, \beta \in (0, 1)$ ; see [Mainik and Schaanning \(2014\)](#) and [Bernardi et al. \(2017\)](#). Our CoVaR and CoES-type allocations with crisis events  $\mathcal{C} = \mathcal{C}^{\text{VaR}_\alpha}$  or  $\mathcal{C}^{\text{ES}_\alpha}$  coincide with those defined in the last displayed equations.

**Remark 1** (Weighted allocations). For a measurable function  $w : \mathbb{R}^d \rightarrow \mathbb{R}_+ := [0, \infty)$ , [Furman and Zitikis \(2008\)](#) proposed the weighted allocation  $q_w(\mathbf{X})$  with the weight function  $w$  being defined by  $q_w(\mathbf{X}) = \mathbb{E}[\mathbf{X}w(\mathbf{X})] / \mathbb{E}[w(\mathbf{X})]$ . By taking an indicator function as weight function  $w(\mathbf{x}) = \mathbf{1}_{\{\mathbf{x} \in \mathcal{C}\}}$  and provided that  $\mathbb{P}(\mathbf{X} \in \mathcal{C}) > 0$ , the weighted allocation coincides with the risk contribution-type systemic allocation  $A^{\mathbb{E}, \dots, \mathbb{E}, \mathcal{C}}$ .

## 2.2. Monte Carlo Estimation of Systemic Risk Allocations

Even if the joint distribution  $F_{\mathbf{X}}$  of the loss random vector  $\mathbf{X}$  is known, the conditional distribution of  $\mathbf{X}$  given  $\mathbf{X} \in \mathcal{C}$ , denoted by  $F_{\mathbf{X}|\mathcal{C}}$ , is typically too complicated to analytically calculate the systemic risk allocations  $A^{q_1, \dots, q_d, \mathcal{C}}$ . An alternative approach is to numerically estimate them by the MC method, as is done in [Yamai and Yoshida \(2002\)](#) and [Fan et al. \(2012\)](#). To this end, assume that one can generate i.i.d. samples from  $F_{\mathbf{X}}$ . If  $\mathbb{P}(\mathbf{X} \in \mathcal{C}) > 0$ , the MC estimator of  $A_j^{q_j, \mathcal{C}}$ ,  $j = 1, \dots, d$  is constructed as follows:

- (1) *Sample from  $\mathbf{X}$* : For a sample size  $N \in \mathbb{N}$ , generate  $\mathbf{X}^{(1)}, \dots, \mathbf{X}^{(N)} \stackrel{\text{ind.}}{\sim} F_{\mathbf{X}}$ .
- (2) *Estimate the crisis event*: If the crisis event  $\mathcal{C}$  contains unknown quantities, replace them with their estimates based on  $\mathbf{X}^{(1)}, \dots, \mathbf{X}^{(N)}$ . Denote by  $\hat{\mathcal{C}}$  the estimated crisis event.
- (3) *Sample from the conditional distribution of  $\mathbf{X}$  given  $\hat{\mathcal{C}}$* : Among  $\mathbf{X}^{(1)}, \dots, \mathbf{X}^{(N)}$ , determine  $\tilde{\mathbf{X}}^{(n)}$  such that  $\tilde{\mathbf{X}}^{(n)} \in \hat{\mathcal{C}}$  for all  $n = 1, \dots, N$ .
- (4) *Construct the MC estimator*: The MC estimate of  $A_j^{q_j, \mathcal{C}}$  is  $q_j(\hat{F}_{\tilde{\mathbf{X}}})$  where  $\hat{F}_{\tilde{\mathbf{X}}}$  is the empirical cdf (ecdf) of the  $\tilde{\mathbf{X}}^{(n)}$ 's.

For an example of (2), if the crisis event is  $\mathcal{C}^{\text{RVaR}_{\alpha_1, \alpha_2}} = \{\mathbf{x} \in \mathbb{R}^d \mid \text{VaR}_{\alpha_1}(S) \leq \mathbf{1}_d^\top \mathbf{x} \leq \text{VaR}_{\alpha_2}(S)\}$ , then  $\text{VaR}_{\alpha_1}(S)$  and  $\text{VaR}_{\alpha_2}(S)$  are unknown parameters, and thus they are replaced by  $\text{VaR}_{\alpha_1}(\hat{F}_S)$  and

$\text{VaR}_{\alpha_2}(\hat{F}_S)$ , where  $\hat{F}_S$  is the ecdf of the total loss  $S^{(n)} := X_1^{(n)} + \dots + X_d^{(n)}$  for  $n = 1, \dots, N$ . By the *law of large numbers (LLN)* and the *central limit theorem (CLT)*, the MC estimator of  $A^{\varrho_1, \dots, \varrho_d, \mathcal{C}}$  is consistent, and the approximate confidence interval of the true allocation can be constructed based on the asymptotic normality; see [Glasserman \(2005\)](#).

The MC cannot handle VaR crisis events if  $S$  admits a pdf, since  $\mathbb{P}(X \in \mathcal{C}^{\text{VaR}_\alpha}) = \mathbb{P}(S = \text{VaR}_\alpha(S)) = 0$ , and thus, no subsample is picked in (3) above. A possible remedy (although the resulting estimator suffers from an inevitable bias) is to replace  $\mathcal{C}^{\text{VaR}}$  with  $\mathcal{C}^{\text{RVaR}_{\alpha-\delta, \alpha+\delta}}$  for sufficiently small  $\delta > 0$ , so that  $\mathbb{P}(S \in \mathcal{C}^{\text{RVaR}_{\alpha-\delta, \alpha+\delta}}) = 2\delta > 0$ .

The main advantage of MC for estimating systemic risk allocations  $A^{\varrho_1, \dots, \varrho_d, \mathcal{C}}$  is that only a random number generator for  $F_X$  is required for implementing the method. Furthermore, MC is applicable for any choice of the crisis event  $\mathcal{C}$  as long as  $\mathbb{P}(X \in \mathcal{C}) > 0$ . Moreover, the main computational load is simulating  $F_X$  in (1) above, which is typically not demanding. The disadvantage of the MC method is its inefficiency concerning the rare-event characteristics of  $\varrho_1, \dots, \varrho_d$  and  $\mathcal{C}$ . To see this, consider the case where  $\mathcal{C} = \mathcal{C}^{\text{RVaR}_{\alpha_1, \alpha_2}}$  and  $\varrho_j = \text{RVaR}_{\beta_1, \beta_2}$  for  $\alpha_1 = \beta_1 = 0.95$  and  $\alpha_2 = \beta_2 = 0.975$ . If the MC sample size is  $N = 10^5$ , there are  $N \times (\alpha_2 - \alpha_1) = 2500$  subsamples resulting from (3). To estimate  $\text{RVaR}_{\beta_1, \beta_2}$  in (4) based on this subsample, only  $2500 \times (\beta_2 - \beta_1) = 62.5$  samples contribute to computing the estimate, which is generally not enough for statistical inference. This effect of sample size reduction is relaxed if ES and/or ES crisis events are considered, but is more problematic for the VaR crisis event since there is a trade-off concerning reducing bias and MC error when choosing  $\delta$ ; see [Koike and Minami \(2019\)](#).

### 3. MCMC Estimation of Systemic Risk Allocations

To overcome the drawback of the MC method for estimating systemic risk allocations, we introduce MCMC methods, which simulate a given distribution by constructing a Markov chain whose stationary distribution is  $F_{X|\mathcal{C}}$ . In this section, we first briefly review MCMC methods, including the MH algorithm as a major subclass of MCMC methods, and then study how to construct an efficient MCMC estimator for the different choices of crisis events.

#### 3.1. A Brief Review of MCMC

Let  $E \subseteq \mathbb{R}^d$  be a set and  $\mathcal{E}$  be a  $\sigma$ -algebra on  $E$ . A *Markov chain* is a sequence of  $E$ -valued random variables  $(\mathbf{X}^{(n)})_{n \in \mathbb{N}_0}$  satisfying the Markov property  $\mathbb{P}(\mathbf{X}^{(n+1)} \in A \mid \mathbf{X}^{(k)} = \mathbf{x}^{(k)}, k \leq n) = \mathbb{P}(\mathbf{X}^{(n+1)} \in A \mid \mathbf{X}^{(n)} = \mathbf{x}^{(n)})$  for all  $n \geq 1$ ,  $A \in \mathcal{E}$ , and  $\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(n)} \in E$ . A Markov chain is characterized by its *stochastic kernel*  $K : E \times \mathcal{E} \rightarrow [0, 1]$  given by  $\mathbf{x} \times A \mapsto K(\mathbf{x}, A) := \mathbb{P}(\mathbf{X}^{(n+1)} \in A \mid \mathbf{X}^{(n)} = \mathbf{x})$ . A probability distribution  $\pi$  satisfying  $\pi(A) = \int_E \pi(d\mathbf{x})K(\mathbf{x}, A)$  for any  $\mathbf{x} \in E$  and  $A \in \mathcal{E}$  is called *stationary distribution*. Assuming  $K(\mathbf{x}, \cdot)$  has a density  $k(\mathbf{x}, \cdot)$ , the *detailed balance condition* (also known as the *reversibility*) with respect to  $\pi$  is given by

$$\pi(\mathbf{x})k(\mathbf{x}, \mathbf{y}) = \pi(\mathbf{y})k(\mathbf{y}, \mathbf{x}), \quad \mathbf{x}, \mathbf{y} \in E, \tag{3}$$

and is known as a sufficient condition for the corresponding kernel  $K$  to have the stationary distribution  $\pi$ ; see [Chib and Greenberg \(1995\)](#). MCMC methods simulate a distribution as a sample path of a Markov chain whose stationary distribution  $\pi$  is the desired one. For a given distribution  $\pi$ , also known as *target distribution*, and a functional  $\varrho$ , the quantity of interest  $\varrho(\pi)$  is estimated by the MCMC estimator  $\varrho(\hat{\pi})$  where  $\hat{\pi}$  is the empirical distribution constructed from a sample path  $\mathbf{X}^{(1)}, \dots, \mathbf{X}^{(N)}$  of the Markov chain whose stationary distribution is  $\pi$ . Under regularity conditions, the MCMC estimator is consistent and asymptotically normal; see [Nummelin \(2002\)](#), [Nummelin \(2004\)](#), and [Meyn and Tweedie \(2012\)](#). Its asymptotic variance can be estimated from  $(\mathbf{X}^{(1)}, \dots, \mathbf{X}^{(N)})$  by, for instance, the *batch means estimator*; see [Jones et al. \(2006\)](#), [Geyer \(2011\)](#) and [Vats et al. \(2015\)](#) for more details. Consequently, one can construct approximate confidence intervals for the true quantity  $\varrho(\pi)$  based on a sample path of the Markov chain.

Since the target distribution  $\pi$  is determined by the problem at hand, the problem is to find the stochastic kernel  $K$  having  $\pi$  as the stationary distribution such that the corresponding Markov chain can be easily simulated. One of the most prevalent stochastic kernels is the *Metropolis–Hastings (MH) kernel*, defined by  $K(x, d\mathbf{y}) = k(x, \mathbf{y})d\mathbf{y} + r(x)\delta_x(d\mathbf{y})$ , where  $\delta_x$  is the Dirac delta function;  $k(x, \mathbf{y}) = q(x, \mathbf{y})\alpha(x, \mathbf{y})$ ;  $q : E \times E \rightarrow \mathbb{R}_+$  is a function called a *proposal density* such that  $x \mapsto q(x, \mathbf{y})$  is measurable for any  $\mathbf{y} \in E$  and  $\mathbf{y} \mapsto q(x, \mathbf{y})$  is a probability density for any  $x \in E$ ;

$$\alpha(x, \mathbf{y}) = \begin{cases} \min \left\{ \frac{\pi(\mathbf{y})q(\mathbf{y}, x)}{\pi(x)q(x, \mathbf{y})}, 1 \right\}, & \text{if } \pi(x)q(x, \mathbf{y}) > 0, \\ 0, & \text{otherwise;} \end{cases}$$

and  $r(x) = 1 - \int_E k(x, \mathbf{y})d\mathbf{y}$ . It can be shown that the MH kernel has stationary distribution  $\pi$ ; see [Tierney \(1994\)](#). Simulation of the Markov chain with this MH kernel is conducted by the *MH algorithm* given in [Algorithm 1](#).

---

**Algorithm 1** Metropolis–Hastings (MH) algorithm.

---

**Require:** Random number generator of the proposal density  $q(x, \cdot)$  for all  $x \in E$ ,  $x^{(0)} \in \text{supp}(\pi)$  and the ratio  $\pi(\mathbf{y})/\pi(x)$  for  $x, \mathbf{y} \in E$ , where  $\pi$  is the density of the stationary distribution.

**Input:** Sample size  $N \in \mathbb{N}$ , proposal density  $q$ , and initial value  $\mathbf{X}^{(0)} = x^{(0)}$ .

**Output:** Sample path  $\mathbf{X}^{(1)}, \dots, \mathbf{X}^{(N)}$  of the Markov chain.

**for**  $n := 0, \dots, N - 1$  **do**

(1) Generate  $\tilde{\mathbf{X}}^{(n)} \sim q(\mathbf{X}^{(n)}, \cdot)$ .

(2) Calculate the *acceptance probability*

$$\alpha_n := \alpha(\mathbf{X}^{(n)}, \tilde{\mathbf{X}}^{(n)}) = \min \left\{ \frac{\pi(\tilde{\mathbf{X}}^{(n)})q(\tilde{\mathbf{X}}^{(n)}, \mathbf{X}^{(n)})}{\pi(\mathbf{X}^{(n)})q(\mathbf{X}^{(n)}, \tilde{\mathbf{X}}^{(n)})}, 1 \right\}. \quad (4)$$

(3) Generate  $U \sim \mathcal{U}(0, 1)$  and set  $\mathbf{X}^{(n+1)} := 1_{[U \leq \alpha_n]} \tilde{\mathbf{X}}^{(n)} + 1_{[U > \alpha_n]} \mathbf{X}^{(n)}$ .

**end for**

---

An advantage of the MCMC method is that a wide variety of distributions can be simulated as a sample path of a Markov chain even if generating i.i.d. samples is not directly feasible. The price to pay is an additional computational cost to calculate the acceptance probability (4), and a possibly higher standard deviation of the estimator  $\varrho(\hat{\pi})$  compared to the standard deviation of estimators constructed from i.i.d. samples. This attributes to the serial dependence among MCMC samples, which can be seen as follows. Suppose first that the candidate  $\tilde{\mathbf{X}}^{(n)}$  is rejected (so  $\{U > \alpha_n\}$  occurs). Then  $\mathbf{X}^{(n+1)} = \mathbf{X}^{(n)}$ , and thus, the samples are perfectly dependent. The candidate  $\tilde{\mathbf{X}}^{(n)}$  is more likely to be accepted if the acceptance probability  $\alpha_n$  is close to 1. In this case,  $\pi(\mathbf{X}^{(n)})$  and  $\pi(\tilde{\mathbf{X}}^{(n)})$  are expected to be close to each other (otherwise,  $\pi(\tilde{\mathbf{X}}^{(n)})/\pi(\mathbf{X}^{(n)})$  and thus  $\alpha_n$  can be small). Under the continuity of  $\pi$ ,  $\tilde{\mathbf{X}}^{(n)}$  and  $\mathbf{X}^{(n)}$  are expected to be close and thus dependent with each other. An efficient MCMC method is such that the candidate  $\tilde{\mathbf{X}}^{(n)}$  is sufficiently far from  $\mathbf{X}^{(n)}$  with the probability  $\pi(\tilde{\mathbf{X}}^{(n)})$  being as close to  $\pi(\mathbf{X}^{(n)})$  as possible. The efficiency of MCMC can indirectly be inspected through the *acceptance rate (ACR)* and the *autocorrelation plot (ACP)*; ACR is the percentage of times a candidate  $\tilde{\mathbf{X}}$  is accepted among the  $N$  iterations, and ACP is the plot of the autocorrelation function of the generated sample path. An efficient MCMC method shows high ACR and steady decline in ACP; see [Chib and Greenberg \(1995\)](#) and [Rosenthal et al. \(2011\)](#) for details. Ideally, the proposal density  $q$  is constructed only based on  $\pi$ , but typically,  $q$  is chosen among a parametric family of distributions. For such cases, simplicity of the choice of tuning parameters of  $q$  is also important.

### 3.2. MCMC Formulation for Estimating Systemic Risk Allocations

Numerous choices of proposal densities  $q$  are possible to construct an MH kernel. In this subsection, we consider how to construct an efficient MCMC method for estimating systemic risk allocations  $A^{q_1, \dots, q_d, \mathcal{C}}$  depending on the choice of the crisis event  $\mathcal{C}$ . Our goal is to directly simulate the conditional distribution  $\mathbf{X}|\mathcal{C}$  by constructing a Markov chain whose stationary distribution is

$$\pi(\mathbf{x}) = f_{\mathbf{X}|\mathbf{X} \in \mathcal{C}}(\mathbf{x}) = \frac{f_{\mathbf{X}}(\mathbf{x})}{\mathbb{P}(\mathbf{X} \in \mathcal{C})} \mathbf{1}_{\{\mathbf{x} \in \mathcal{C}\}}, \quad \mathbf{x} \in E \subseteq \mathbb{R}^d, \tag{5}$$

provided  $\mathbb{P}(\mathbf{X} \in \mathcal{C}) > 0$ . Samples from this distribution can directly be used to estimate systemic risk allocations with crisis event  $\mathcal{C}$  and arbitrary marginal risk measures  $q_1, \dots, q_d$ . Other potential applications are outlined in Remark 2.

**Remark 2** (Gini shortfall allocation). *Samples from the conditional distribution  $F_{\mathbf{X}|\mathcal{C}_\alpha^{\text{ES}}}$  can be used to estimate, for example, the tail-Gini coefficient  $\text{TGini}_\alpha(X_j, S) = \frac{4}{1-\alpha} \text{Cov}(X_j, F_S(S) | S \geq \text{VaR}_\alpha(S))$  for  $\alpha \in (0, 1)$ , and the Gini shortfall allocation (Furman et al. (2017))  $\text{GS}_\alpha(X_j, S) = \mathbb{E}[X_j | S \geq \text{VaR}_\alpha(S)] + \lambda \cdot \text{TGini}_\alpha(X_j, S)$ ,  $\lambda \in \mathbb{R}_+$  more efficiently than by applying the MC method. Another application is to estimate risk allocations derived by optimization, given a constant economic capital; see Laeven and Govaerts (2004) and Dhaene et al. (2012).*

We now construct an MH algorithm with target distribution (5). To this end, we assume that

1. the ratio  $f_{\mathbf{X}}(\mathbf{y})/f_{\mathbf{X}}(\mathbf{x})$  can be evaluated for any  $\mathbf{x}, \mathbf{y} \in \mathcal{C}$ , and that
2. the support of  $f_{\mathbf{X}}$  is  $\mathbb{R}^d$  or  $\mathbb{R}_+^d$ .

Regarding Assumption 1, the normalization constant of  $f_{\mathbf{X}}$  and the probability  $\mathbb{P}(\mathbf{X} \in \mathcal{C})$  are not necessary to be known, since they cancel out in the numerator and the denominator of  $\pi(\mathbf{y})/\pi(\mathbf{x})$ . In Assumption 2, the loss random vector  $\mathbf{X}$  refers to the profit&loss (P&L) if  $\text{supp}(f_{\mathbf{X}}) = \mathbb{R}^d$ , and to pure losses if  $\text{supp}(f_{\mathbf{X}}) = \mathbb{R}_+^d$ . Note that the case  $\text{supp}(f_{\mathbf{X}}) = [c_1, \infty) \times \dots \times [c_d, \infty)$ ,  $c_1, \dots, c_d \in \mathbb{R}$  is essentially included in the case of pure losses as long as the marginal risk measures  $q_1, \dots, q_d$  are law invariant and translation invariant, and the crisis event is the set of linear constraints of Form (2). To see this, define  $\tilde{X}_j = X_j - c_j$ ,  $j = 1, \dots, d$ ,  $\tilde{\mathbf{X}} = (\tilde{X}_1, \dots, \tilde{X}_d)$  and  $\mathbf{c} = (c_1, \dots, c_d)$ . Then  $\text{supp}(f_{\tilde{\mathbf{X}}}) = \mathbb{R}_+^d$  and  $\mathbf{X}|\mathcal{C} \stackrel{d}{=} \tilde{\mathbf{X}}|\tilde{\mathcal{C}} + \mathbf{c}$ , where  $\tilde{\mathcal{C}}$  is the set of linear constraints with parameters  $\tilde{h}_m = h_m$  and  $\tilde{v}_m = v_m - \mathbf{h}_m^\top \mathbf{c}$ . By law invariance and translation invariance of  $q_1, \dots, q_d$ ,

$$q_j(X_j|\mathbf{X} \in \mathcal{C}) = c_j + q_j(\tilde{X}_j|\tilde{\mathbf{X}} \in \tilde{\mathcal{C}}), \quad j = 1, \dots, d.$$

Therefore, the problem of estimating  $A^{q_1, \dots, q_d, \mathcal{C}}(\mathbf{X})$  reduces to that of estimating  $A^{q_1, \dots, q_d, \tilde{\mathcal{C}}}(\tilde{\mathbf{X}})$  for the shifted loss random vector  $\tilde{\mathbf{X}}$  (such that  $\text{supp}(f_{\tilde{\mathbf{X}}}) = \mathbb{R}_+^d$ ) and the modified crisis event of the same form.

For the P&L case, the RVaR and ES crisis events are the set of linear constraints of Form (2) with the number of constraints  $M = 2$  and  $1$ , respectively. In the case of pure losses, additional  $d$  constraints  $\mathbf{e}_{j,d}^\top \mathbf{x} \geq 0$ ,  $j = 1, \dots, d$  are imposed, where  $\mathbf{e}_{j,d}$  is the  $j$ th  $d$ -dimensional unit vector. Therefore, the RVaR and ES crisis events are of Form (2) with  $M = d + 2$  and  $d + 1$ , respectively. For the VaR crisis event,  $\mathbb{P}(\mathbf{X} \in \mathcal{C}) = 0$ , and thus, (5) cannot be properly defined. In this case, the allocation  $A^{q_1, \dots, q_d, \mathcal{C}^{\text{VaR}}}$  depends on the conditional joint distribution  $\mathbf{X}|\mathcal{C}_\alpha^{\text{VaR}}$ , but is completely determined by its first  $d' := d - 1$  variables  $(X_1, \dots, X_{d'})|\mathcal{C}_\alpha^{\text{VaR}}$ , since  $X_d|\mathcal{C}_\alpha^{\text{VaR}} \stackrel{d}{=} (\text{VaR}_\alpha(S) - \sum_{j=1}^{d'} X_j)|\mathcal{C}_\alpha^{\text{VaR}} \stackrel{d}{=} \text{VaR}_\alpha(S) - \sum_{j=1}^{d'} X_j|\mathcal{C}_\alpha^{\text{VaR}}$ . Estimating systemic risk allocations under the VaR crisis event can thus be achieved by simulating the target distribution



$$\begin{aligned} \pi^{\text{VaR}_\alpha}(\mathbf{x}') &= f_{\mathbf{X}'|S=\text{VaR}_\alpha(S)}(\mathbf{x}) = \frac{f_{(\mathbf{X}',S)}(\mathbf{x}', \text{VaR}_\alpha(S))}{f_S(\text{VaR}_\alpha(S))} \\ &= \frac{f_{\mathbf{X}}(\mathbf{x}', \text{VaR}_\alpha(S) - \mathbf{1}_{d'}^\top \mathbf{x}')}{f_S(\text{VaR}_\alpha(S))} \mathbf{1}_{[\text{VaR}_\alpha(S) - \mathbf{1}_{d'}^\top \mathbf{x}' \in \text{supp}(f_d)]}, \quad \mathbf{x}' \in \mathbb{R}^{d'}, \end{aligned} \tag{6}$$

where  $\mathbf{X}' = (X_1, \dots, X_{d'})$  and the last equation is derived from the linear transformation  $(\mathbf{X}', S) \mapsto \mathbf{X}$  with unit Jacobian. Note that other transformations are also possible; see [Betancourt \(2012\)](#). Under Assumption 1, the ratio  $\pi^{\text{VaR}_\alpha}(\mathbf{y}) / \pi^{\text{VaR}_\alpha}(\mathbf{x})$  can be evaluated and  $f_S(\text{VaR}_\alpha(S))$  is not required to be known. In the case of pure losses, the target distribution  $\pi^{\text{VaR}_\alpha}$  is subject to  $d$  linear constraints  $e_{j,d'}^\top \mathbf{x}' \geq 0, j = 1, \dots, d'$  and  $\mathbf{1}_{d'}^\top \mathbf{x}' \geq \text{VaR}_\alpha(S)$ , where the first  $d'$  constraints come from the non-negativity of the losses and the last one is from the indicator in (6). Therefore, the crisis event  $\mathcal{C}^{\text{VaR}}$  for  $(X_1, \dots, X_{d'})$  is of Form (2). In the case of P&L,  $\text{supp}(f_d) = \mathbb{R}$  and  $\text{VaR}_\alpha(S) - \mathbf{1}_{d'}^\top \mathbf{x}' \in \text{supp}(f_d)$  holds for any  $\mathbf{x}' \in \mathbb{R}^{d'}$ . Therefore, the target distribution (6) is free from any constraints and the problem reduces to constructing an MCMC method with target distribution  $\pi(\mathbf{x}') \propto f_{\mathbf{X}}(\mathbf{x}', \text{VaR}_\alpha(S) - \mathbf{1}_{d'}^\top \mathbf{x}')$ ,  $\mathbf{x}' \in \mathbb{R}^{d'}$ . In this paper, the P&L case with VaR crisis event is not investigated further, since our focus is the simulation of constrained target distributions; see [Koike and Minami \(2019\)](#) for an MCMC estimation in the P&L case.

MCMC methods to simulate constrained target distributions require careful design of the proposal density  $q$ . A simple MCMC method is *Metropolis–Hastings with rejection* in which the support of the proposal density  $q$  may not coincide with that of the target distribution, which is the crisis event  $\mathcal{C}$ , and a candidate is immediately rejected when it violates the constraints. This construction of MCMC is often inefficient due to a low acceptance probability, especially around the boundary of  $\mathcal{C}$ . In this case, an efficient MCMC method can be expected only when the probability mass of  $\pi$  is concentrated near the center of  $\mathcal{C}$ . In the following sections, we introduce two alternative MCMC methods for the constrained target distributions  $F_{\mathbf{X}|\mathcal{C}}$  of interest, the HMC method and the GS. Each of them is applicable and can be efficient for different choices of the crisis event and underlying loss distribution functions  $F_{\mathbf{X}}$ .

### 3.3. Estimation with Hamiltonian Monte Carlo

We find that if the HMC method is applicable, it is typically the most preferable method to simulate constrained target distributions because of its efficiency and ease of handling constraints. In Section 3.3.1, we briefly present the HMC method with a reflection for constructing a Markov chain supported on the constrained space. In Section 3.3.2, we propose a heuristic for determining the parameters of the HMC method based on the MC presamples.

#### 3.3.1. Hamiltonian Monte Carlo with Reflection

For the possibly unnormalized target density  $\pi$ , consider the *potential energy*  $U(\mathbf{x})$ , *kinetic energy*  $K(\mathbf{p})$ , and the *Hamiltonian*  $H(\mathbf{x}, \mathbf{p})$  defined by

$$U(\mathbf{x}) = -\log \pi(\mathbf{x}), \quad K(\mathbf{p}) = -\log f_K(\mathbf{p}) \quad \text{and} \quad H(\mathbf{x}, \mathbf{p}) = U(\mathbf{x}) + K(\mathbf{p}),$$

with *position variable*  $\mathbf{x} \in E$ , *momentum variable*  $\mathbf{p} \in \mathbb{R}^d$ , and *kinetic energy density*  $f_K(\mathbf{p})$  such that  $f_K(-\mathbf{p}) = f_K(\mathbf{p})$ . In this paper, the kinetic energy distribution  $F_K$  is set to be the multivariate standard normal with  $K(\mathbf{p}) = \frac{1}{2} \mathbf{p}^\top \mathbf{p}$  and  $\nabla K(\mathbf{p}) = \mathbf{p}$ ; other choices of  $F_K$  are discussed in Appendix B.2. In the HMC method, a Markov chain augmented on the state space  $E \times \mathbb{R}^d$  with the stationary distribution  $\pi(\mathbf{x})f_K(\mathbf{p})$  is constructed and the desired samples from  $\pi$  are obtained as the first  $|E|$ -dimensional margins. A process  $(\mathbf{x}(t), \mathbf{p}(t)), t \in \mathbb{R}$  on  $E \times \mathbb{R}^d$  is said to follow the *Hamiltonian dynamics* if it follows the ordinary differential equation (ODE)

$$\frac{d}{dt} \mathbf{x}(t) = \nabla K(\mathbf{p}), \quad \frac{d}{dt} \mathbf{p}(t) = -\nabla U(\mathbf{x}). \tag{7}$$

Through the Hamiltonian dynamics, the Hamiltonian  $H$  and the volume are conserved, that is,  $dH(\mathbf{x}(t), \mathbf{p}(t))/dt = 0$  and the map  $(\mathbf{x}(0), \mathbf{p}(0)) \mapsto (\mathbf{x}(t), \mathbf{p}(t))$  has a unit Jacobian for any  $t \in \mathbb{R}$ ; see Neal et al. (2011). Therefore, the value of the joint target density  $\pi \cdot f_K$  remains unchanged by the Hamiltonian dynamics, that is,

$$\pi(\mathbf{x}(0))f_K(\mathbf{p}(0)) = \exp(-H(\mathbf{x}(0), \mathbf{p}(0))) = \exp(-H(\mathbf{x}(t), \mathbf{p}(t))) = \pi(\mathbf{x}(t))f_K(\mathbf{p}(t)), \quad t \geq 0.$$

In practice, the dynamics (7) are discretized for simulation by, for example, the so-called *leapfrog method* summarized in Algorithm 2; see Leimkuhler and Reich (2004) for other discretization methods.

---

**Algorithm 2** Leapfrog method for Hamiltonian dynamics.

---

**Input:** Current states  $(\mathbf{x}(0), \mathbf{p}(0))$ , stepsize  $\epsilon > 0$ , gradients  $\nabla U$  and  $\nabla K$ .

**Output:** Updated position  $(\mathbf{x}(\epsilon), \mathbf{p}(\epsilon))$ .

- (1)  $\mathbf{p}(\frac{\epsilon}{2}) = \mathbf{p}(0) - \frac{\epsilon}{2} \nabla U(\mathbf{x}(0))$ .
  - (2)  $\mathbf{x}(\epsilon) = \mathbf{x}(0) + \epsilon \nabla K(\mathbf{p}(\frac{\epsilon}{2}))$ .
  - (3)  $\mathbf{p}(\epsilon) = \mathbf{p}(\epsilon/2) + \frac{\epsilon}{2} \nabla U(\mathbf{x}(\epsilon))$ .
- 

Note that the evaluation of  $\nabla U$  does not require the normalization constant of  $\pi$  to be known, since  $\nabla U = -(\nabla \pi) / \pi$ . By repeating the leapfrog method  $T$  times with stepsize  $\epsilon$ , the Hamiltonian dynamics are approximately simulated with length  $T\epsilon$ . Due to the discretization error, the Hamiltonian is not exactly preserved, while it is expected to be almost preserved for  $\epsilon$  which is small enough. The discretization error  $H(\mathbf{x}(T\epsilon), \mathbf{p}(T\epsilon)) - H(\mathbf{x}(0), \mathbf{p}(0))$  is called the *Hamiltonian error*.

All the steps of the HMC method are described in Algorithm 3. In Step (1), the momentum variable is first updated from  $\mathbf{p}(0)$  to  $\mathbf{p}$ , where  $\mathbf{p}$  follows the kinetic energy distribution  $F_K$  so that the value of the Hamiltonian  $H = -\log(\pi \cdot f_K)$  changes. In Step (3), the current state  $(\mathbf{x}(0), \mathbf{p})$  is moved along the level curve of  $H(\mathbf{x}(0), \mathbf{p})$  by simulating the Hamiltonian dynamics.

---

**Algorithm 3** Hamiltonian Monte Carlo to simulate  $\pi$ .

---

**Require:** Random number generator of  $F_K$ ,  $\mathbf{x}^{(0)} \in \text{supp}(\pi)$ ,  $\pi(\mathbf{y}) / \pi(\mathbf{x})$ ,  $\mathbf{x}, \mathbf{y} \in E$  and  $f_K(\mathbf{p}') / f_K(\mathbf{p})$ ,  $\mathbf{p}, \mathbf{p}' \in \mathbb{R}^d$ .

**Input:** Sample size  $N \in \mathbb{N}$ , kinetic energy density  $f_K$ , target density  $\pi$ , gradients of the potential and kinetic energies  $\nabla U$  and  $\nabla K$ , stepsize  $\epsilon > 0$ , integration time  $T \in \mathbb{N}$  and initial position  $\mathbf{X}^{(0)} = \mathbf{x}^{(0)}$ .

**Output:** Sample path  $\mathbf{X}^{(1)}, \dots, \mathbf{X}^{(N)}$  of the Markov chain.

**for**  $n := 0, \dots, N - 1$  **do**

- (1) Generate  $\mathbf{p}^{(n)} \sim F_K$ .
- (2) Set  $(\tilde{\mathbf{X}}^{(n)}, \tilde{\mathbf{p}}^{(n)}) = (\mathbf{X}^{(n)}, \mathbf{p}^{(n)})$ .
- (3) **for**  $t := 1, \dots, T$ ,

$$(\tilde{\mathbf{X}}^{(n+t/T)}, \tilde{\mathbf{p}}^{(n+t/T)}) = \text{Leapfrog}(\tilde{\mathbf{X}}^{(n+(t-1)/T)}, \tilde{\mathbf{p}}^{(n+(t-1)/T)}, \epsilon, \nabla U, \nabla K).$$

**end for**

- (4)  $\tilde{\mathbf{p}}^{(n+1)} = -\mathbf{p}^{(n+1)}$ .
- (5) Calculate  $\alpha_n = \min \left\{ \frac{\pi(\tilde{\mathbf{X}}^{(n+1)})f_K(\tilde{\mathbf{p}}^{(n+1)})}{\pi(\mathbf{X}^{(n)})f_K(\mathbf{p}^{(n)})}, 1 \right\}$ .
- (6) Set  $\mathbf{X}^{(n+1)} := 1_{[U \leq \alpha_n]} \tilde{\mathbf{X}}^{(n+1)} + 1_{[U > \alpha_n]} \mathbf{X}^{(n)}$  for  $U \sim U(0, 1)$ .

**end for**

---

By flipping the momentum in Step (4), the HMC method is shown to be reversible w.r.t.  $\pi$  (c.f. (3)) and thus to have the stationary distribution  $\pi$ ; see Neal et al. (2011) for details. Furthermore, by the conservation property of the Hamiltonian dynamics, the acceptance probability in Step (5) is expected to be close to 1. Moreover, by taking  $T$  as sufficiently large, the candidate  $\tilde{X}^{(n+1)}$  is expected to be sufficiently decorrelated from the current position  $X^{(n)}$ . Consequently, the resulting Markov chain is expected to be efficient.

The remaining challenge for applying the HMC method to our problem of estimating systemic risk allocations is how to handle the constraint  $\mathcal{C}$ . As we have seen in Sections 2.1 and 3.2,  $\mathcal{C}$  is assumed to be an intersection of linear constraints with parameters  $(h_m, v_m)$ ,  $m = 1, \dots, M$  describing hyperplanes. Following the ordinary leapfrog method, a candidate is immediately rejected when the trajectory of the Hamiltonian dynamics penetrates one of these hyperplanes. To avoid it, we modify the leapfrog method according to the reflection technique introduced in Afshar and Domke (2015) and Chevallier et al. (2018). As a result, the trajectory is reflected when it hits a hyperplane and the Markov chain moves within the constrained space with probability one. Details of the HMC method with the reflection for our application are described in Appendix A.

### 3.3.2. Choice of Parameters for HMC

HMC requires as input two parameters, the *stepsize*  $\epsilon$ , and the *integration time*  $T$ . As we now explain, neither of them should be chosen too large nor too small. Since the stepsize  $\epsilon$  controls the accuracy of the simulation of the Hamiltonian dynamics,  $\epsilon$  needs to be small enough to approximately conserve the Hamiltonian; otherwise, the acceptance probability can be much smaller than 1. On the other hand, an  $\epsilon$  which is too small requires the integration time  $T$  to be large enough for the trajectory to reach a farther distance, which is computationally costly. Next, the integration time  $T$  needs to be large enough to decorrelate the candidate state with the current state. Meanwhile, the trajectory of the Hamiltonian dynamics may make a U-turn and come back to the starting point if the integration time  $T$  is too long; see Neal et al. (2011) for an illustration of this phenomenon.

A notable characteristic of our problem of estimating systemic risk allocations is that the MC sample from the target distribution  $\pi$  is available but its sample size may not be sufficient for statistical inference, and, in the case of the VaR crisis event, the samples only approximately follow the target distribution. We utilize the information of this *MC presample* to build a heuristic for determining the parameters  $(\epsilon, T)$ ; see Algorithm 4.

In this heuristic, the initial stepsize is set to be  $\epsilon = c_\epsilon d^{-1/4}$  for some constant  $c_\epsilon > 0$ , say,  $c_\epsilon = 1$ . This scale was derived in Beskos et al. (2010) and Beskos et al. (2013) under certain assumptions on the target distribution. We determine  $\epsilon$  through the relationship with the acceptance probability. In Step (2-2-2-1) of Algorithm 4, multiple trajectories are simulated, starting from each MC presample with the current stepsize  $\epsilon$ . In the next Step (2-2-2-2), we monitor the acceptance probability and the distance between the starting and ending points while extending the trajectories. Based on the asymptotic optimal acceptance probability 0.65 (c.f. Gupta et al. (1990) and Betancourt et al. (2014)) as  $d \rightarrow \infty$ , we set the *target acceptance probability* as

$$\underline{\alpha} = \frac{1 + (d - 1) \times 0.65}{d} \in (0.65, 1].$$

The stepsize is gradually decreased in Step (2-1) of Algorithm 4 until the minimum acceptance probability calculated in Step (2-3) exceeds  $\underline{\alpha}$ . To prevent the trajectory from a U-turn, in Step (2-2-2-3), each trajectory is immediately stopped when the distance begins to decrease. The resulting integration time is set to be the average of these turning points, as seen in Step (3). Note that other termination conditions of extending trajectories are possible; see Hoffman and Gelman (2014) and Betancourt (2016).

---

**Algorithm 4** Heuristic for determining the stepsize  $\epsilon$  and integration time  $T$ .

---

**Input:** MC presample  $\mathbf{X}_1^{(0)}, \dots, \mathbf{X}_{N_0}^{(0)}$ , gradients  $\nabla U$  and  $\nabla K$ , target acceptance probability  $\underline{\alpha}$ , initial constant  $c_\epsilon > 0$  and the maximum integration time  $T_{\max}$  ( $c_\epsilon = 1$  and  $T_{\max} = 1000$  are set as default values).

**Output:** Stepsize  $\epsilon$  and integration time  $T$ .

(1) Set  $\alpha_{\min} = 0$  and  $\epsilon = c_\epsilon d^{-1/4}$ .

(2) **while**  $\alpha_{\min} < \underline{\alpha}$

(2-1) Set  $\epsilon = \epsilon/2$ .

(2-2) **for**  $n := 1, \dots, N_0$

(2-2-1) Generate  $\mathbf{p}_n^{(0)} \sim F_K$ .

(2-2-2) **for**  $t := 1, \dots, T_{\max}$

(2-2-2-1) Set  $\mathbf{Z}_n^{(t)} = \text{Leapfrog}(\mathbf{Z}_n^{(t-1)}, \epsilon, \nabla U, \nabla K)$  for  $\mathbf{Z}_n^{(t-1)} = (\mathbf{X}_n^{(t-1)}, \mathbf{p}_n^{(t-1)})$ .

(2-2-2-2) Calculate

$$\alpha_{n,t} = \alpha(\mathbf{Z}_n^{(t-1)}, \mathbf{Z}_n^{(t)}) \quad \text{and} \quad \Delta_t = \|\mathbf{X}_n^{(t)} - \mathbf{X}_n^{(0)}\| - \|\mathbf{X}_n^{(t-1)} - \mathbf{X}_n^{(0)}\|.$$

(2-2-2-3) **if**  $\Delta_t < 0$  and  $\Delta_{t-1} > 0$ , **break** and set  $T_n^* = t - 1$ .

**end for**

**end for**

(2-3) Compute  $\alpha_{\min} = \min(\alpha_{n,t} \mid t = 1, 2, \dots, T_n^*, n = 1, \dots, N_0)$ .

**end while**

(3) Set  $T = \lfloor \frac{1}{N_0} \sum_{n=1}^{N_0} T_n^* \rfloor$ .

---

At the end of this section, we briefly revisit the choice of the kinetic energy distribution  $F_K$ , which is taken to be a multivariate standard normal throughout this work. As discussed in Neal et al. (2011), applying the HMC method with target distribution  $\pi$  and kinetic energy distribution  $N(\mathbf{0}, \Sigma^{-1})$  is equivalent to applying HMC with the standardized target distribution  $\mathbf{x} \rightarrow \pi(L\mathbf{x})$  and  $F_K = N(\mathbf{0}, \mathbf{I})$ , where  $L$  is the Cholesky factor of  $\Sigma$  such that  $\Sigma = LL^\top$ . By taking  $\Sigma$  to be the covariance matrix of  $\pi$ , the standardized target distribution becomes uncorrelated with unit variances. In our problem, the sample covariance matrix  $\hat{\Sigma} = \hat{L}\hat{L}^\top$  calculated based on the MC presample is used alternatively. The new target distribution  $\tilde{\pi}(\mathbf{y}) = \pi(\hat{L}\mathbf{y})|\hat{L}|$  where  $|\hat{L}|$  denotes the Jacobian of  $\hat{L}$ , is almost uncorrelated with unit variances, and thus the standard normal kinetic energy fits well; see Livingstone et al. (2019). If the crisis event consists of the set of linear constraints  $(\mathbf{h}_m, v_m)$ ,  $m = 1, \dots, M$ , then the standardized target density is also subject to the set of linear constraints  $(\hat{L}^\top \mathbf{h}_m, v_m)$ ,  $m = 1, \dots, M$ . Since the ratio  $f_X(\hat{L}\mathbf{y})/f_X(\hat{L}\mathbf{x})$  can still be evaluated under Assumption 1, we conclude that the problem remains unchanged after standardization.

Theoretical results of the HMC method with normal kinetic energy are available only when  $\mathcal{C}$  is bounded (Cances et al. (2007) and Chevallier et al. (2018)), or when  $\mathcal{C}$  is unbounded and the tail of  $\pi$  is roughly as light as that of the normal distribution (Livingstone et al. (2016) and Durmus et al. (2017)). Boundedness of  $\mathcal{C}$  holds for VaR and RVaR crisis events with pure losses; see Koike and Minami (2019). As is discussed in this paper, convergence results of MCMC estimators are accessible when the density of the underlying joint loss distribution is bounded from above on  $\mathcal{C}$ , which is typically the case when the underlying copula does not admit lower tail dependence. For other cases where  $\mathcal{C}$  is unbounded or the density explodes on  $\mathcal{C}$ , no convergence results are available. Potential remedies for the HMC method to deal with heavy-tailed target distributions are discussed in Appendix B.2.

### 3.4. Estimation with Gibbs Sampler

As discussed in Section 3.3.2, applying HMC methods to heavy-tailed target distributions on unbounded crisis events is not theoretically supported. To deal with this case, we introduce the GS in this section.

#### 3.4.1. True Gibbs Sampler for Estimating Systemic Risk Allocations

The GS is a special case of the MH method in which the proposal density  $q$  is completely determined by the target density  $\pi$  via

$$q_{GS}(\mathbf{x}, \mathbf{y}) = \sum_{i=(i_1, \dots, i_d) \in \mathcal{I}_d} p_i \pi(y_{i_1} | \mathbf{x}_{-i_1}) \pi(y_{i_2} | y_{i_1}, \mathbf{x}_{-(i_1, i_2)}) \cdots \pi(y_{i_d} | \mathbf{y}_{-i_d}), \tag{8}$$

where  $\mathbf{x}_{-(j_1, \dots, j_l)}$  is the  $(d - l)$ -dimensional vector that excludes the components  $j_1, \dots, j_l$  from  $\mathbf{x}$ ,  $\pi(x_j | \mathbf{x}_{-j}) = \pi_{j|-j}(x_j | \mathbf{x}_{-j})$  is the conditional density of the  $j$ th variable of  $\pi$  given all the other components,  $\mathcal{I}_d \subseteq \{1, \dots, d\}^d$  is the so-called *index set*, and  $(p_i \in [0, 1], i \in \mathcal{I}_d)$  is the *index probability distribution* such that  $\sum_{i \in \mathcal{I}_d} p_i = 1$ . For this choice of  $q$ , the acceptance probability is always equal to 1; see Johnson (2009). The GS is called *deterministic scan (DSGS)* if  $\mathcal{I}_d = \{(1, \dots, d)\}$  and  $p_{(1, \dots, d)} = 1$ . When the index set is the set of permutations of  $(1, \dots, d)$ , the GS is called *random permutation (RPGS)*. Finally, the *random scan GS (RSGS)* has the proposal (8) with  $\mathcal{I}_d = \{1, \dots, d\}^d$  and  $p_{(i_1, \dots, i_d)} = p_{i_1} \cdots p_{i_d}$  with probabilities  $(p_1, \dots, p_d) \in (0, 1)^d$  such that  $\sum_{j=1}^d p_j = 1$ . These three GSs can be shown to have  $\pi$  as stationary distribution; see Johnson (2009).

Provided that the *full conditional distributions*  $\pi_{j|-j}, j = 1, \dots, d$  can be simulated, the proposal distribution (8) can be simulated by first selecting an index  $i \in \mathcal{I}_d$  with probability  $p_i$  and then replacing the  $j$ th component of the current state with a sample from  $\pi_{j|-j}$  sequentially for  $j = i_1, \dots, i_d$ . The main advantage of the GS is that the tails of  $\pi$  are naturally incorporated via full conditional distributions, and thus the MCMC method is expected to be efficient even if  $\pi$  is heavy-tailed. On the other hand, the applicability of the GS is limited to target distributions such that  $\pi_{j|-j}$  is available. Moreover, fast simulators of  $\pi_{j|-j}, j = 1, \dots, d$ , are required, since the computational time linearly increases w.r.t. the dimension  $d$ .

In our problem of estimating systemic risk allocations, we find that the GS is applicable when the crisis event is of the form

$$\mathcal{C} = \{\mathbf{x} \in \mathbb{R}^d \text{ or } \mathbb{R}_+^d \mid v_1 \leq \mathbf{h}^\top \mathbf{x} \leq v_2\}, \quad v_1, v_2 \in \mathbb{R} \cup \{\pm\infty\}, \quad \mathbf{h} = (h_1, \dots, h_d) \in \mathbb{R}^d \setminus \{\mathbf{0}_d\}. \tag{9}$$

The RVaR crisis event is obviously a special case of (9), and the ES crisis event is included as a limiting case for  $v_2 \rightarrow \infty$ . Furthermore, the full conditional copulas of the underlying joint loss distribution and their inverses are required to be known as we now explain. Consider the target density  $\pi = f_{\mathbf{X} | v_1 \leq \mathbf{h}^\top \mathbf{X} \leq v_2}$ . For its  $j$ th full conditional density  $\pi_{j|-j}(x_j | \mathbf{x}_{-j})$ , notice that

$$\{v_1 \leq \mathbf{h}^\top \mathbf{X} \leq v_2, \mathbf{X}_{-j} = \mathbf{x}_{-j}\} = \left\{ \frac{v_1 - \mathbf{h}_{-j}^\top \mathbf{x}_{-j}}{h_j} \leq X_j \leq \frac{v_2 - \mathbf{h}_{-j}^\top \mathbf{x}_{-j}}{h_j}, \mathbf{X}_{-j} = \mathbf{x}_{-j} \right\}$$

and thus, for  $v_{i,j}(\mathbf{x}_{-j}) = (v_i - \mathbf{h}_{-j}^\top \mathbf{x}_{-j}) / h_j, i = 1, 2$ , we obtain the cdf of  $\pi_{j|-j}$  as

$$F_{X_j | (v_1 \leq \mathbf{h}^\top \mathbf{X} \leq v_2, \mathbf{X}_{-j} = \mathbf{x}_{-j})}(x_j) = \frac{F_{X_j | \mathbf{X}_{-j} = \mathbf{x}_{-j}}(x_j) - F_{X_j | \mathbf{X}_{-j} = \mathbf{x}_{-j}}(v_{1,j}(\mathbf{x}_{-j}))}{F_{X_j | \mathbf{X}_{-j} = \mathbf{x}_{-j}}(v_{2,j}(\mathbf{x}_{-j})) - F_{X_j | \mathbf{X}_{-j} = \mathbf{x}_{-j}}(v_{1,j}(\mathbf{x}_{-j}))} \tag{10}$$

for  $v_{1,j}(\mathbf{x}_{-j}) \leq x_j \leq v_{2,j}(\mathbf{x}_{-j})$ . Denoting the denominator of (10) by  $\Delta_j(\mathbf{x}_{-j})$ , we obtain the quantile function

$$F_{X_j|v_{1,j} \leq h^\top \mathbf{X} \leq v_{2,j}, \mathbf{X}_{-j}=\mathbf{x}_{-j}}^{-1}(u) = F_{X_j|\mathbf{X}_{-j}=\mathbf{x}_{-j}}^{-1}\left(\Delta_j(\mathbf{x}_{-j}) \cdot u + F_{X_j|\mathbf{X}_{-j}=\mathbf{x}_{-j}}(v_{1,j}(\mathbf{x}_{-j}))\right).$$

Therefore, if  $F_{X_j|\mathbf{X}_{-j}=\mathbf{x}_{-j}}$  and its quantile function are available, one can simulate the full conditional target densities  $\pi_{j|-j}$  with the inversion method; see Devroye (1985). Availability of  $F_{X_j|\mathbf{X}_{-j}=\mathbf{x}_{-j}}$  and its inverse typically depends on the copula of  $\mathbf{X}$ . By Sklar’s theorem (1), the  $j$ th full conditional distribution of  $F_{\mathbf{X}}$  can be written as

$$F_{X_j|\mathbf{X}_{-j}=\mathbf{x}_{-j}}(x_j) = C_{j|-j}(F_j(x_j) | \mathbf{F}_{-j}(\mathbf{x}_{-j})),$$

where  $\mathbf{F}_{(j_1, \dots, j_l)}(\mathbf{x}_{(j_1, \dots, j_l)}) = (F_{j_1}(x_{j_1}), \dots, F_{j_l}(x_{j_l}))$ ,  $-(j_1, \dots, j_l) = \{1, \dots, d\} \setminus \{j_1, \dots, j_l\}$  and  $C_{j|-j}$  is the  $j$ th full conditional copula defined by

$$C_{j|-j}(u_j | \mathbf{u}_{-j}) = \mathbb{P}(U_j \leq u_j | \mathbf{U}_{-j} = \mathbf{u}_{-j}) = \frac{D_{-j}C(\mathbf{u})}{D_{-j}C(u_1, \dots, u_{j-1}, 1, u_{j+1}, \dots, u_d)},$$

where  $D$  denotes the operator of partial derivatives with respect to the components given as subscripts and  $\mathbf{U} \sim C$ . Assuming the full conditional copula  $C_{j|-j}$  and its inverse  $C_{j|-j}^{-1}$  are available, one can simulate  $\tilde{X}_j \sim \pi_{j|-j}$  via

$$\begin{aligned} U &\sim U(0, 1), \\ \tilde{U} &= U + (1 - U)C_{j|-j}(F_j(v_1(\mathbf{x}_{-j}) | \mathbf{F}_{-j}(\mathbf{x}_{-j})), \\ \tilde{X}_j &= F_j^{-1} \circ C_{j|-j}^{-1}(\tilde{U} | \mathbf{F}_{-j}(\mathbf{x}_{-j})). \end{aligned}$$

Examples of copulas for which the full conditional distributions and their inverses are available include normal, Student  $t$ , and Clayton copulas; see Cambou et al. (2017). In this case, the GS is also applicable to the corresponding survival ( $\pi$ -rotated) copula  $\hat{C}$ , since

$$\hat{C}_{j|-j}(\mathbf{u}) = 1 - C_{j|-j}(1 - u_j | \mathbf{1}_{d'} - \mathbf{u}_{-j}), \quad \hat{C}_{j|-j}^{-1}(\mathbf{u}) = 1 - C_{j|-j}^{-1}(1 - u_j | \mathbf{1}_{d'} - \mathbf{u}_{-j}), \quad j = 1, \dots, d,$$

by the relationship  $\tilde{\mathbf{U}} = \mathbf{1} - \mathbf{U} \sim \hat{C}$  for  $\mathbf{U} \sim C$ . In a similar way, one can also obtain full conditional copulas and their inverses for other rotated copulas; see Hofert et al. (2018) Section 3.4.1 for rotated copulas.

In the end, we remark that even if the full conditional distributions and their inverses are not available,  $\pi_{j|-j}$  can be simulated by, for example, the acceptance and rejection method, or even the MH algorithm; see Appendix B.3.

### 3.4.2. Choice of Parameters for GS

As discussed in Section 3.3.2, we use information from the MC presamples to determine the parameters of the Gibbs kernel (8). Note that standardization of the variables as applied in the HMC method in Section 3.3.2 is not available for the GS, since the latter changes the underlying joint loss distribution, and since the copula after rotating variables is generally not accessible, except for in the elliptical case; see Christen et al. (2017). Among the presented variants of GSs, we adopt RSGS, since determining  $d$  probabilities  $(p_1, \dots, p_d)$  is relatively easy, whereas RPGS requires  $d!$  probabilities to be determined. To this end, we consider the RSGS with the parameters  $(p_1, \dots, p_d)$  determined by a heuristic described in Algorithm 5.

---

**Algorithm 5** Random scan Gibbs sampler (RSGS) with heuristic to determine  $(p_1, \dots, p_d)$ .

---

**Require:** Random number generator of  $\pi_{j|-j}$  and  $\mathbf{x}^{(0)} \in \text{supp}(\pi)$ .

**Input:** MC presample  $\tilde{\mathbf{X}}_1^{(0)}, \dots, \tilde{\mathbf{X}}_{N_0}^{(0)}$ , sample size  $N \in \mathbb{N}$ , initial state  $\mathbf{x}^{(0)}$ , sample size of the pre-run  $N_{\text{pre}}$  and the target autocorrelation  $\rho$  ( $N_{\text{pre}} = 100$  and  $\rho = 0.15$  are set as default values).

**Output:**  $N$  sample path  $\mathbf{X}^{(1)}, \dots, \mathbf{X}^{(N)}$  of the Markov chain.

(1) Compute the sample covariance matrix  $\hat{\Sigma}$  based on  $\tilde{\mathbf{X}}_1^{(0)}, \dots, \tilde{\mathbf{X}}_{N_0}^{(0)}$ .

(2) Set  $p_j \propto \hat{\Sigma}_{j,j} - \hat{\Sigma}_{j,-j} \hat{\Sigma}_{-j,-j}^{-1} \hat{\Sigma}_{-j,j}$  and  $\mathbf{X}^{(0)} = \mathbf{X}_{\text{pre}}^{(0)} = \mathbf{x}^{(0)}$ .

(3) **for**  $n := 1, \dots, N_{\text{pre}}$

(3-1) Generate  $J = j$  with probability  $p_j$ .

(3-2) Update  $X_{\text{pre},J}^{(n)} \sim \pi_{J|-J}(\cdot | \mathbf{X}_{\text{pre}}^{(n-1)})$  and  $\mathbf{X}_{\text{pre},-J}^{(n)} = \mathbf{X}_{\text{pre},-J}^{(n-1)}$ .

**end for**

(4) Set

$$T = \operatorname{argmin}_{h \in \mathbb{N}_0} \left\{ \text{estimated autocorrelations of } \mathbf{X}_{\text{pre}}^{(1)}, \dots, \mathbf{X}_{\text{pre}}^{(N_{\text{pre}})} \text{ with lag } h \leq \rho \right\}.$$

(5) **for**  $n := 1, \dots, N, t := 1, \dots, T$

(5-1) Generate  $J = j$  with probability  $p_j$ .

(5-2) Update  $X_J^{(n-1+t/T)} \sim \pi_{J|-J}(\cdot | \mathbf{X}^{(n-1+(t-1)/T)})$  and  $\mathbf{X}_{-J}^{(n-1+t/T)} = \mathbf{X}_{-J}^{(n-1+(t-1)/T)}$ .

**end for**

---

The RSGS kernel is simulated in Steps (3) and (5) of Algorithm 5. To determine the selection probabilities  $p_1, \dots, p_d$ , consider a one-step update of the RSGS from  $\mathbf{X}^{(n)}$  to  $\mathbf{X}^{(n+1)}$  with  $\mathbf{X}^{(n)} \sim \pi$  and the one-step kernel

$$K_{\text{RSGS}}(\mathbf{x}, \mathbf{y}) = \sum_{j=1}^d p_j \pi_{j|-j}(y_j | \mathbf{x}_{-j}) \mathbf{1}_{[y_{-j} = \mathbf{x}_{-j}]}.$$

Liu et al. (1995, Lemma 3) implies that

$$\operatorname{Cov}(X_j^{(n)}, X_j^{(n+1)}) = \sum_{i=1}^d p_i \mathbb{E}[\mathbb{E}[X_j | \mathbf{X}_{-i}]] = \sum_{i=1}^d p_i \{m_j^{(2)} - \mathbb{E}[\operatorname{Var}(X_j | \mathbf{X}_{-i})]\} \propto - \sum_{i=1}^d p_i \mathbb{E}[\operatorname{Var}(X_j | \mathbf{X}_{-i})],$$

where  $m_j^{(k)}$  is the  $k$ th moment of  $\pi_j$ .

For the objective function  $\sum_{j=1}^d \operatorname{Cov}(X_j^{(n)}, X_j^{(n+1)})$ , its minimizer  $(p_1^*, \dots, p_d^*)$  under the constraint  $\sum_{j=1}^d p_j = 1$  satisfies

$$p_j^* \propto \mathbb{E}[\operatorname{Var}(X_j | \mathbf{X}_{-j})]. \tag{11}$$

While this optimizer can be computed based on the MC presamples, we observed that its stable estimation is as computationally demanding as estimating the risk allocations themselves. Alternatively, we calculate (11) under the assumption that  $\pi$  follows an elliptical distribution. Under this assumption, (11) is given by

$$p_j \propto \Sigma_{j,j} - \Sigma_{j,-j} \Sigma_{-j,-j}^{-1} \Sigma_{-j,j},$$

where  $\Sigma$  is the covariance matrix of  $\pi$  and  $\Sigma_{J_1, J_2}, J_1, J_2 \subseteq \{1, \dots, d\}$  is the submatrix of  $\Sigma$  with indices in  $J_1 \times J_2$ . As seen in Step (2) of Algorithm 5,  $\Sigma$  is replaced by its estimate based on the MC presamples.

As shown in Christen et al. (2017), Gibbs samplers require a large number of iterations to lower the serial correlation when the target distribution has strong dependence. To reduce serial correlations,

we take every  $T$ th sample in Step (5-2), where  $T \in \mathbb{N}$  is called the *thinning interval of times*. Note that we use the same notation  $T$  as that of the integration time in HMC, since they both represent a repetition time of some single step. Based on the preliminary run with length  $N_{\text{pre}}$  in Step (3) in Algorithm 5,  $T$  is determined as the smallest lag  $h$  such that the marginal autocorrelations with lag  $h$  are all smaller than the target autocorrelation  $\rho$ ; see Step (4) in Algorithm 5.

#### 4. Numerical Experiments

In this section, we demonstrate the performance of the MCMC methods for estimating systemic risk allocations by a series of numerical experiments. We first conduct a simulation study in which true allocations or their partial information are available. Then, we perform an empirical study to demonstrate that our MCMC methods are applicable to a more practical setup. Finally, we make more detailed comparisons between the MC and MCMC methods in various setups. All experiments were run on a MacBook Air with 1.4 GHz Intel Core i5 processor and 4 GB 1600 MHz of DDR3 RAM.

##### 4.1. Simulation Study

In this simulation study, we compare the estimates and standard errors of the MC and MCMC methods under the low-dimensional risk models described in Section 4.1.1. The results and discussions are summarized in Section 4.1.2.

##### 4.1.1. Model Description

We consider the following three-dimensional loss distributions:

- (M1) *generalized Pareto distributions (GPDs)* with parameters  $(\xi_j, \beta_j) = (0.3, 1)$  and survival Clayton copula with parameter  $\theta = 2$  so that Kendall's tau equals  $\tau = \theta / (\theta + 2) = 0.5$ ;
- (M2) *multivariate Student  $t$  distribution* with  $\nu = 5$  degrees of freedom, location vector  $\mathbf{0}$ , and dispersion matrix  $\Sigma = (\rho_{i,j})$ , where  $\rho_{j,j} = 1$  and  $\rho_{i,j} = |i - j|/d$  for  $i, j = 1, \dots, d, i \neq j$ .

Since the marginals are homogeneous and the copula is exchangeable, the systemic risk allocations under the loss distribution (M1) are all equal, provided that the crisis event is invariant under the permutation of the variables. For the loss distribution (M2), by ellipticality of the joint distribution, analytical formulas of risk contribution-type systemic risk allocations are available; see McNeil et al. (2015) Corollary 8.43. The parameters of the distributions (M1) and (M2) take into account the stylized facts that the loss distribution is heavy-tailed and extreme losses are positively dependent.

We consider the VaR, RVaR, and ES crisis events with confidence levels  $\alpha^{\text{VaR}} = 0.99$ ,  $(\alpha_1^{\text{RVaR}}, \alpha_2^{\text{RVaR}}) = (0.975, 0.99)$  and  $\alpha^{\text{ES}} = 0.99$ , respectively. For each crisis event, the risk contribution, VaR, RVaR, and ES-type systemic risk allocations are estimated by the MC and MCMC methods, where the parameters of the marginal risk measures VaR, RVaR, and ES are set to be  $\beta^{\text{VaR}} = 0.99$ ,  $(\beta_1^{\text{RVaR}}, \beta_2^{\text{RVaR}}) = (0.975, 0.99)$  and  $\beta^{\text{ES}} = 0.99$ , respectively.

We first conduct the MC simulation for the distributions (M1) and (M2). For the VaR crisis event, the modified event  $\mathcal{C}^{\text{mod}} = \{\text{VaR}_{\alpha-\delta}(S) \leq \mathbf{1}_d^\top \mathbf{x} \leq \text{VaR}_{\alpha+\delta}(S)\}$  with  $\delta = 0.001$  is used to ensure that  $\mathbb{P}(\mathbf{X} \in \mathcal{C}^{\text{mod}}) > 0$ . Based on these MC presamples, the Markov chains are constructed as described in Sections 3.3 and 3.4. For the MCMC method, (M1) is the case of pure losses and (M2) is the case of P&L. Therefore, the HMC method is applied to the distribution (M1) for the VaR and RVaR crisis events, the GS is applied to (M1) for the ES crisis event and the GS is applied to the distribution (M2) for the RVaR and ES crisis events. The target distribution of (M2) with VaR constraint is free from constraints and was already investigated in Koike and Minami (2019); we thus omit this case and consider the five remaining cases.

Note that 99.8% of the MC samples from the unconditional distribution are discarded for the VaR crisis event and a further 97.5% of them are wasted to estimate the RVaR contributions. Therefore,  $1/(0.002 \times 0.025) = 10^5/5 = 20,000$  MC samples are required to obtain one MC sample from the



conditional distribution. Taking this into account, the sample size of the MC estimator is set to be  $N_{MC} = 10^5$ . The sample size of the MCMC estimators is free from such constraints and thus is chosen to be  $N_{MCMC} = 10^4$ . Initial values  $x_0$  for the MCMC methods are taken as the mean vector calculated from the MC samples. Biases are computed only for the contribution-type allocations in the distribution (M2) since the true values are available in this case. For all the five cases, the MC and the MCMC standard errors are computed according to Glasserman (2013) Chapter 1, for MC, and Jones et al. (2006) for MCMC. Asymptotic variances of the MCMC estimators are estimated by the batch means estimator with batch length  $L_N := \lceil N^{\frac{1}{2}} \rceil = 100$  and batch size  $B_N := \lceil N/L_N \rceil = 100$ . The results are summarized in Tables 1 and 2.

**Table 1.** Estimates and standard errors for the MC and HMC estimators of risk contribution, RVaR, VaR, and ES-type systemic risk allocations under (I) the VaR crisis event, and (II) the RVaR crisis event for the loss distribution (M1). The sample size of the MC method is  $N_{MC} = 10^5$ , and that of the HMC method is  $N_{MCMC} = 10^4$ . The acceptance rate (ACR), stepsize  $\epsilon$ , integration time  $T$ , and run time are  $ACR = 0.996$ ,  $\epsilon = 0.210$ ,  $T = 12$ , and run time = 1.277 mins in Case (I), and  $ACR = 0.984$ ,  $\epsilon = 0.095$ ,  $T = 13$ , and run time = 1.649 mins in Case (II).

Estimator	MC			HMC		
	$A_1^{e,C}(X)$	$A_2^{e,C}(X)$	$A_3^{e,C}(X)$	$A_1^{e,C}(X)$	$A_2^{e,C}(X)$	$A_3^{e,C}(X)$
(I) GPD + survival Clayton with VaR crisis event: $\{S = VaR_{0.99}(S)\}$						
$\mathbb{E}[X C^{VaR}]$	9.581	9.400	9.829	9.593	9.599	9.619
<b>Standard error</b>	0.126	0.118	0.120	0.007	0.009	0.009
$RVaR_{0.975,0.99}(X C^{VaR})$	12.986	12.919	13.630	13.298	13.204	13.338
<b>Standard error</b>	0.229	0.131	0.086	0.061	0.049	0.060
$VaR_{0.99}(X C^{VaR})$	13.592	13.235	13.796	13.742	13.565	13.768
<b>Standard error</b>	0.647	0.333	0.270	0.088	0.070	0.070
$ES_{0.99}(X C^{VaR})$	14.775	13.955	14.568	14.461	14.227	14.427
<b>Standard error</b>	0.660	0.498	0.605	0.192	0.176	0.172
(II) GPD + survival Clayton with RVaR crisis event: $\{VaR_{0.975}(S) \leq S \leq VaR_{0.99}(S)\}$						
$\mathbb{E}[X C^{RVaR}]$	7.873	7.780	7.816	7.812	7.802	7.780
<b>Standard error</b>	0.046	0.046	0.046	0.012	0.012	0.011
$RVaR_{0.975,0.99}(X C^{RVaR})$	11.790	11.908	11.680	11.686	11.696	11.646
<b>Standard error</b>	0.047	0.057	0.043	0.053	0.055	0.058
$RVaR_{0.99}(X C^{VaR})$	12.207	12.382	12.087	12.102	12.053	12.044
<b>Standard error</b>	0.183	0.197	0.182	0.074	0.069	0.069
$ES_{0.99}(X C^{RVaR})$	13.079	13.102	13.059	12.859	12.791	12.713
<b>Standard error</b>	0.182	0.173	0.188	0.231	0.218	0.187

**Table 2.** Estimates and standard errors for the MC and the GS estimators of risk contribution, VaR, RVaR, and ES-type systemic risk allocations under (III) distribution (M1) and the ES crisis event, (IV) distribution (M2), and the RVaR crisis event, and (V) distribution (M2) and ES crisis event. The sample size of the MC method is  $N_{MC} = 10^5$  and that of the GS is  $N_{MCMC} = 10^4$ . The thinning interval of times  $T$ , selection probability  $p$  and run time are  $T = 12$ ,  $p = (0.221, 0.362, 0.416)$  and run time = 107.880 secs in Case (III),  $T = 10$ ,  $p = (0.330, 0.348, 0.321)$  and run time = 56.982 secs in Case (IV) and  $T = 4$ ,  $p = (0.241, 0.503, 0.255)$  and run time = 22.408 secs in Case (V).

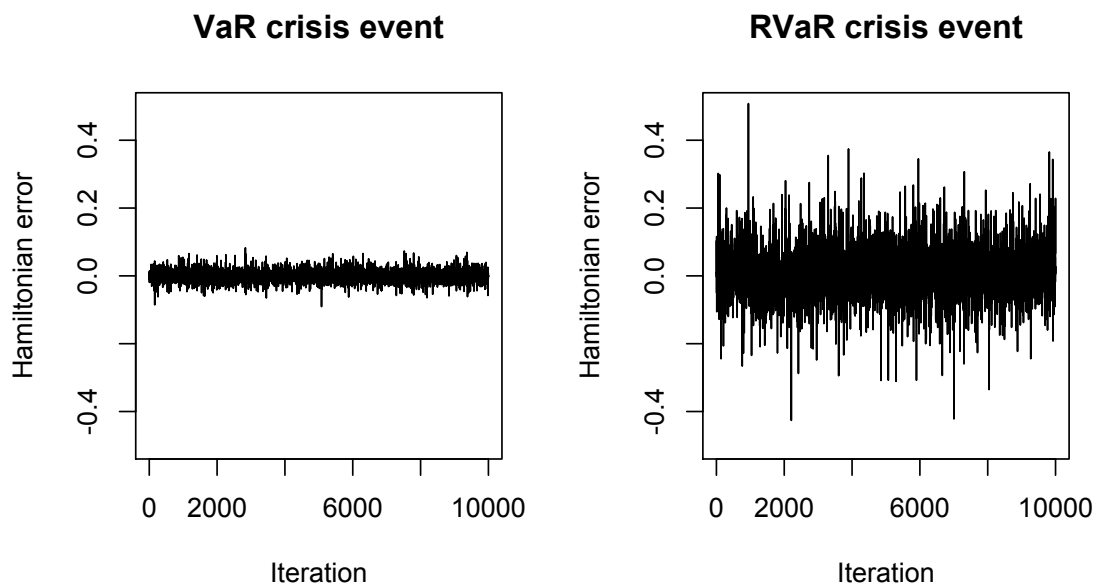
Estimator	MC			GS		
	$A_1^{e,C}(X)$	$A_2^{e,C}(X)$	$A_3^{e,C}(X)$	$A_1^{e,C}(X)$	$A_2^{e,C}(X)$	$A_3^{e,C}(X)$
(III) GPD + survival Clayton with ES crisis event: $\{VaR_{0.99}(S) \leq S\}$						
$E[X C^{ES}]$	15.657	15.806	15.721	15.209	15.175	15.190
<b>Standard error</b>	0.434	0.475	0.395	0.257	0.258	0.261
$RVaR_{0.975,0.99}(X C^{ES})$	41.626	41.026	45.939	45.506	45.008	45.253
<b>Standard error</b>	1.211	1.065	1.615	1.031	1.133	1.256
$VaR_{0.99}(X C^{ES})$	49.689	48.818	57.488	55.033	54.746	54.783
<b>Standard error</b>	4.901	4.388	4.973	8.079	5.630	3.803
$ES_{0.99}(X C^{ES})$	104.761	109.835	97.944	71.874	72.588	70.420
<b>Standard error</b>	23.005	27.895	17.908	4.832	4.584	4.313
(IV) Multivariate $t$ with RVaR crisis event: $\{VaR_{0.975}(S) \leq S \leq VaR_{0.99}(S)\}$						
$E[X C^{RVaR}]$	2.456	1.934	2.476	2.394	2.060	2.435
<b>Bias</b>	0.019	-0.097	0.038	-0.043	0.029	-0.002
<b>Standard error</b>	0.026	0.036	0.027	0.014	0.023	0.019
$RVaR_{0.975,0.99}(X C^{RVaR})$	4.670	4.998	4.893	4.602	5.188	4.748
<b>Standard error</b>	0.037	0.042	0.031	0.032	0.070	0.048
$RVaR_{0.99}(X C^{VaR})$	5.217	5.397	5.240	4.878	5.717	5.092
<b>Standard error</b>	0.238	0.157	0.145	0.049	0.174	0.100
$ES_{0.99}(X C^{RVaR})$	5.929	5.977	5.946	5.446	6.517	6.063
<b>Standard error</b>	0.204	0.179	0.199	0.156	0.248	0.344
(V) Multivariate $t$ with ES crisis event: $\{VaR_{0.99}(S) \leq S\}$						
$E[X C^{ES}]$	3.758	3.099	3.770	3.735	3.126	3.738
<b>Bias</b>	0.017	-0.018	0.029	-0.005	0.009	-0.003
<b>Standard error</b>	0.055	0.072	0.060	0.031	0.027	0.030
$RVaR_{0.975,0.99}(X C^{ES})$	8.516	8.489	9.051	8.586	8.317	8.739
<b>Standard error</b>	0.089	0.167	0.161	0.144	0.156	0.158
$VaR_{0.99}(X C^{ES})$	9.256	9.754	10.327	9.454	9.517	9.890
<b>Standard error</b>	0.517	0.680	0.698	0.248	0.293	0.327
$ES_{0.99}(X C^{ES})$	11.129	12.520	12.946	11.857	12.469	12.375
<b>Standard error</b>	0.595	1.321	0.826	0.785	0.948	0.835

#### 4.1.2. Results and Discussions

Since a fast random number generators are available for the joint loss distributions (M1) and (M2), the MC estimators are computed almost instantly. On the other hand, the MCMC methods cost around 1.5 min for simulating the  $N = 10^4$  MCMC samples, as reported in Tables 1 and 2. For the HMC

method, the main computational cost consists of calculating gradients  $N \times T$  times for the leapfrog method, and calculating the ratio of target densities  $N$  times in the acceptance/rejection step, where  $N$  is the length of the sample path and  $T$  is the integration time. For the GS, simulating an  $N$ -sample path requires  $N \times T \times d$  random numbers from the full conditional distributions, where  $T$  here is the thinning interval of times. Therefore, the computational time of the GS linearly increases w.r.t. the dimension  $d$ , which can become prohibitive for the GS in high dimensions. To save computational time, MCMC methods generally require careful implementations of calculating the gradients and the ratio of the target densities for HMC, and of simulating the full conditional distributions for GS.

Next, we inspect the performance of the HMC and GS methods. We observed that the autocorrelations of all sample paths steadily decreased below 0.1 if lags were larger than 15. Together with the high ACRs, we conclude that the Markov chains can be considered to be converged. According to the heuristic in Algorithm 4, the stepsize and integration time for the HMC method are selected to be  $(\epsilon, T) = (0.210, 12)$  in Case (I) and  $(\epsilon, T) = (0.095, 13)$  in Case (II). As indicated by the small Hamiltonian errors in Figure 1, the acceptance rates in both cases are quite close to 1.



**Figure 1.** Hamiltonian errors of the HMC methods for estimating systemic risk allocations with VaR (left) and RVaR (right) crisis events for the loss distribution (M1). The stepsize and integration time are set to be  $(\epsilon, T) = (0.210, 12)$  in Case (I) and  $(\epsilon, T) = (0.095, 13)$  in Case (II).

For the GS, the thinning interval of times  $T$  and the selection probability  $p$  are determined as  $T = 12$  and  $p = (0.221, 0.362, 0.416)$  in Case (III),  $T = 10$  and  $p = (0.330, 0.348, 0.321)$  in Case (IV) and  $T = 4$  and  $p = (0.241, 0.503, 0.255)$  in Case (V). For biases of the estimators, observe that in all cases ((I) to (V)), the estimates of the MC method and the MCMC method are close to each other. In Cases (I), (II), and (III), the true allocations are the homogeneous allocations, whereas their exact values are not known. From the estimates in Tables 1 and 2, the MCMC estimates are, on average, more equally allocated compared to those of the MC method, especially in Case (III) where heavy-tailedness may lead to quite slow convergence rates of the MC method. Therefore, lower biases of the MCMC estimators are obtained, compared to those of the MC estimators. In the case of risk contributions in Case (IV) and (V), exact biases are computed based on ellipticity, and they show that the GS estimator has a smaller bias than the one of the MC estimator.

Although the MC sample size is 10 times larger than that of the MCMC method, the standard error of the latter is, in most cases, smaller than the MC standard error. This improvement becomes larger as the probability of the crisis event becomes smaller. The largest improvement is observed in Case (I) with the VaR crisis event, and the smallest one is in Cases (III) and (V) with the ES crisis

event. MCMC estimates of the risk contribution-type allocations have consistently smaller standard errors than the MC ones. For the R VaR, VaR, and ES-type allocations, the improvement of standard error varies according to the loss models and the crisis event. A notable improvement is observed for ES-type allocation in Case (III), although a stable statistical inference is challenging due to the heavy-tailedness of the target distribution.

Overall, the simulation study shows that the MCMC estimators outperform the MC estimators due to the increased effective sample size and its insusceptibility to the probability of the crisis event. The MCMC estimators are especially recommended when the probability of the crisis event is too small for the MC method to sufficiently simulate many samples for a meaningful statistical analysis.

**Remark 3** (Joint loss distributions with negative dependence in the tail). *In the above simulation study, we only considered joint loss distributions with positive dependence. Under the existence of positive dependence, the target density  $f_{\mathbf{X}|v_\alpha \leq S \leq v_\beta}$  puts more probability mass around its mean, and the probability decays as the point moves away from the mean, since positive dependence among  $X_1, \dots, X_d$  prevents them from going in opposite directions (i.e., one component increases and another one decreases) under the sum constraint; see Koike and Minami (2019) for details. This phenomenon leads to the target distributions being more centered and elliptical, which in turn facilitates efficient moves of Markov chains. Although it may not be realistic, joint loss distributions with negative dependence in the tail are also possible. In this case, the target distribution has more variance, heavy tails, and is even multimodal, since two components can move in opposite directions under the sum constraint. For such cases, constructing efficient MCMC methods becomes more challenging; see Lan et al. (2014) for a remedy for multimodal target distributions with Riemannian manifold HMC.*

#### 4.2. Empirical Study

In this section, we illustrate our suggested MCMC methods for estimating risk allocations from insurance company indemnity claims. The dataset consists of 1500 liability claims provided by the Insurance Services Office. Each claim contains an indemnity payment  $X_1$  and an allocated loss adjustment expense (ALAE)  $X_2$ ; see Hogg and Klugman (2009) for a description. The joint distribution of losses and expenses is studied, for example, in Frees and Valdez (1998) and Klugman and Parsa (1999). Based on Frees and Valdez (1998), we adopt the following parametric model:

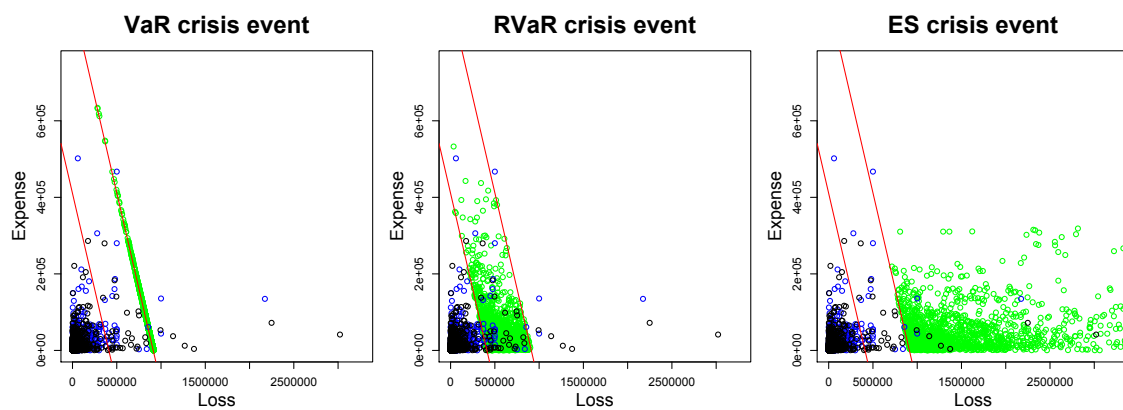
- (M3) univariate marginals are  $X_1 \sim \text{Par}(\lambda_1, \theta_1)$  and  $X_2 \sim \text{Par}(\lambda_2, \theta_2)$  with  $(\lambda_1, \theta_1) = (14, 036, 1.122)$  and  $(\lambda_2, \theta_2) = (14, 219, 2.118)$ , and the copula is the survival Clayton copula with parameter  $\theta = 0.512$  (which corresponds to Spearman's rho  $\rho_S = 0.310$ ).

Note that in the loss distribution (M3), the Gumbel copula used in Frees and Valdez (1998) is replaced by the survival Clayton copula, since both of them have the same type of tail dependence and the latter possesses more computationally tractable derivatives. The parameter of the survival Clayton copula is determined so that it reaches the same Spearman's rho observed in Frees and Valdez (1998). Figure 2 illustrates the data and samples from the distribution (M3). Our goal is to calculate the VaR, R VaR, and ES-type allocations with VaR, R VaR, and ES crisis events for the same confidence levels as in Section 4.1.1. We apply the HMC method to all three crisis events since, due to the infinite and finite variances of  $X_1$  and  $X_2$ , respectively, the optimal selection probability of the second variable calculated in Step 2 of Algorithm 5 is quite close to 0, and thus the GS did not perform well. The simulated HMC samples are illustrated in Figure 2. The results of estimating the systemic risk allocations are summarized in Table 3.

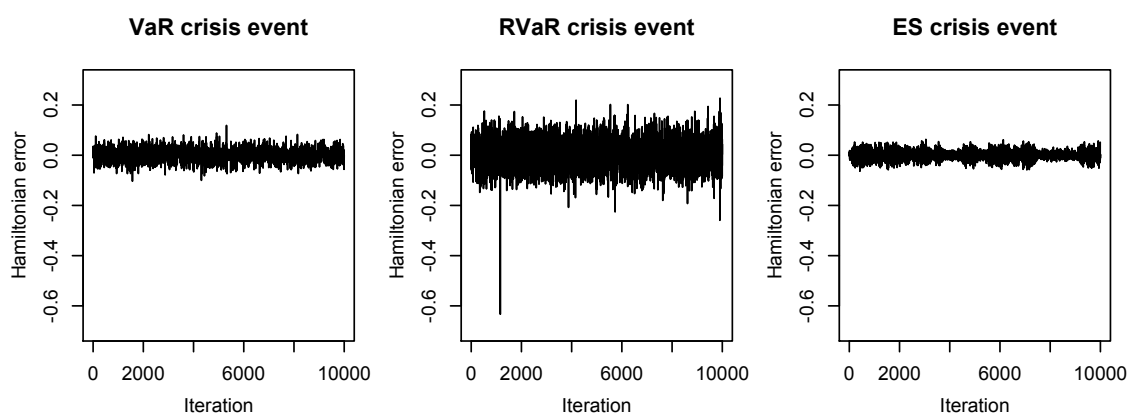
The HMC samples shown in Figure 2 indicate that the conditional distributions of interest are successfully simulated from the desired regions. As displayed in Figure 3, the Hamiltonian errors of all three HMC methods are sufficiently small, which led to the high ACRs of 0.997, 0.986, and 0.995, as listed in Table 3. We also observed that autocorrelations of all sample paths steadily decreased below 0.1 if lags were larger than 80. Together with the high ACRs, we conclude that the Markov chains can

be considered to be converged. Due to the heavy-tailedness of the target distribution in the case of the ES crisis event, the stepsize is very small and the integration time is very large compared to the former two cases of the VaR and RVaR crisis events. As a result, the HMC algorithm in this case has a long run time.

The estimates of the MC and HMC methods are close in all cases, except Case (III). In Case (III), the HMC estimates are smaller than the MC ones in almost all cases. Based on the much smaller standard errors of HMC, one could infer that the MC estimates are likely overestimating the allocations due to a small number of extremely large losses, although the corresponding conditional distribution is extremely heavy-tailed, and thus no estimation method might be reliable. In terms of the standard error, the estimation of systemic risk allocations by the HMC method were improved in Cases (I) and (III) compared to that of the MC method; the MC standard errors are slightly smaller than those of HMC in Case (II). All results considered, we conclude from this empirical study that the MCMC estimators outperform the MC estimators in terms of standard error. On the other hand, as indicated by the theory of HMC with normal kinetic energy, the HMC method is not recommended for heavy-tailed target distributions due to the long computational time caused by a small stepsize and large integration time determined by Algorithm 5.



**Figure 2.** Plots of  $N = 1500$  MCMC samples (green) with VaR (left), RVaR (center), and ES (right) crisis events. All plots include the data and the MC samples with sample size  $N = 1500$  in black and blue dots, respectively. The red lines represent  $x_1 + x_2 = \widehat{\text{VaR}}_{\alpha_1}(S)$  and  $x_1 + x_2 = \widehat{\text{VaR}}_{\alpha_2}(S)$  where  $\widehat{\text{VaR}}_{\alpha_1}(S) = 4.102 \times 10^4$  and  $\widehat{\text{VaR}}_{\alpha_2}(S) = 9.117 \times 10^4$  are the MC estimates of  $\text{VaR}_{\alpha_1}(S)$  and  $\text{VaR}_{\alpha_2}(S)$ , respectively, for  $\alpha_1 = 0.975$  and  $\alpha_2 = 0.99$ .



**Figure 3.** Hamiltonian errors of the HMC methods for estimating systemic risk allocations with VaR, RVaR, and ES crisis events for the loss distribution (M3). The stepsize and the integration time are chosen as  $(\epsilon, T) = (0.015, 34)$ ,  $(\epsilon, T) = (0.026, 39)$  and  $(\epsilon, T) = (5.132 \times 10^{-5}, 838)$ , respectively.

**Table 3.** Estimates and standard errors for the MC and HMC estimators of RVaR, VaR, and ES-type systemic risk allocations under the loss distribution (M3) with the (I) VaR crisis event, (II) RVaR crisis event, and (III) ES crisis event. The MC sample size is  $N_{MC} = 10^5$ , and that of the HMC method is  $N_{MCMC} = 10^4$ . The acceptance rate (ACR), stepsize  $\epsilon$ , integration time  $T$ , and run time are ACR = 0.997,  $\epsilon = 0.015$ ,  $T = 34$  and run time = 2.007 min in Case (I), ACR = 0.986,  $\epsilon = 0.026$ ,  $T = 39$  and run time = 2.689 min in Case (II), ACR = 0.995,  $\epsilon = 5.132 \times 10^{-5}$ ,  $T = 838$  and run time = 44.831 min in Case (III).

Estimator	MC		HMC	
	$A_1^{e,C}(X)$	$A_2^{e,C}(X)$	$A_1^{e,C}(X)$	$A_2^{e,C}(X)$
(I) VaR crisis event: $\{S = \text{VaR}_{0.99}(S)\}$				
$\mathbb{E}[X C^{\text{VaR}}]$	842465.497	73553.738	844819.901	71199.334
<b>Standard error</b>	7994.573	7254.567	6306.836	6306.836
$\text{RVaR}_{0.975,0.99}(X C^{\text{VaR}})$	989245.360	443181.466	915098.833	428249.307
<b>Standard error</b>	307.858	24105.163	72.568	20482.914
$\text{VaR}_{0.99}(X C^{\text{VaR}})$	989765.514	500663.072	915534.362	615801.118
<b>Standard error</b>	4670.966	54576.957	669.853	96600.963
$\text{ES}_{0.99}(X C^{\text{VaR}})$	990839.359	590093.887	915767.076	761038.843
<b>Standard error</b>	679.055	75024.692	47.744	31211.908
(II) RVaR crisis event: $\{\text{VaR}_{0.975}(S) \leq S \leq \text{VaR}_{0.99}(S)\}$				
$\mathbb{E}[X C^{\text{RVaR}}]$	528455.729	60441.368	527612.751	60211.561
<b>Standard error</b>	3978.477	2119.461	4032.475	2995.992
$\text{RVaR}_{0.975,0.99}(X C^{\text{RVaR}})$	846956.570	349871.745	854461.670	370931.946
<b>Standard error</b>	1866.133	6285.523	2570.997	9766.697
$\text{VaR}_{0.99}(X C^{\text{RVaR}})$	865603.369	413767.829	871533.550	437344.509
<b>Standard error</b>	5995.341	29105.059	12780.741	21142.135
$\text{ES}_{0.99}(X C^{\text{RVaR}})$	882464.968	504962.099	885406.811	529034.580
<b>Standard error</b>	3061.110	17346.207	3134.144	23617.278
(III) ES crisis event: $\{\text{VaR}_{0.99}(S) \leq S\}$				
$\mathbb{E}[X C^{\text{ES}}]$	8663863.925	137671.653	2934205.458	140035.782
<b>Standard error</b>	3265049.590	10120.557	165794.772	14601.958
$\text{RVaR}_{0.975,0.99}(X C^{\text{ES}})$	35238914.131	907669.462	17432351.450	589309.196
<b>Standard error</b>	2892208.689	31983.660	443288.649	3471.641
$\text{VaR}_{0.99}(X C^{\text{ES}})$	56612082.905	1131248.055	20578728.307	615572.940
<b>Standard error</b>	1353975.612	119460.411	1364899.752	12691.776
$\text{ES}_{0.99}(X C^{\text{ES}})$	503537848.192	2331984.181	25393466.446	649486.810
<b>Standard error</b>	268007317.199	468491.127	1138243.137	7497.200

### 4.3. Detailed Comparison of MCMC with MC

In the previous numerical experiments, we fixed the dimensions of the portfolios and confidence levels of the crisis events. Comparing the MC and MCMC methods after balancing against computational time might be more reasonable, although one should keep in mind that run time depends on various external factors, such as the implementation, hardware, workload, programming language, or compiler options (and our implementation was not optimized for any of these factors). In this section, we compare the MC and MCMC methods with different dimensions, confidence levels,

and parameters of the HMC methods in terms of bias, standard error, and the mean squared error (MSE), adjusted by run time.

In this experiment, we fix the sample size of the MC and MCMC methods as  $N_{MC} = N_{MCMC} = 10^4$ . In addition, we assume  $X \sim t_\nu(\mathbf{0}, P)$ , that is, the joint loss follows the multivariate Student  $t$  distribution with  $\nu = 6$  degrees of freedom, location vector  $\mathbf{0}$ , and dispersion matrix  $P$ , which is the correlation matrix with all off-diagonal entries equal to  $1/12$ . The dimension  $d$  of the loss portfolio will vary for comparison. We consider only risk contribution-type systemic risk allocations under VaR, RVaR, and ES crisis events, as true values of these allocations are available to compare against; see [McNeil et al. \(2015\)](#), Corollary 8.43. If  $b$  and  $\sigma$  denote the bias and standard deviation of the MC or MCMC estimator and  $S$  the run time, then (under the assumption that run time linearly increases by sample size) we define the *time-adjusted MSEs* by

$$MSE_{MC} = b_{MC}^2 + \frac{\sigma_{MC}^2}{\frac{S_{MCMC}}{S_{MC}} \times N_{MCMC}} \quad \text{and} \quad MSE_{MCMC} = b_{MCMC}^2 + \frac{\sigma_{MCMC}^2}{N_{MCMC}}.$$

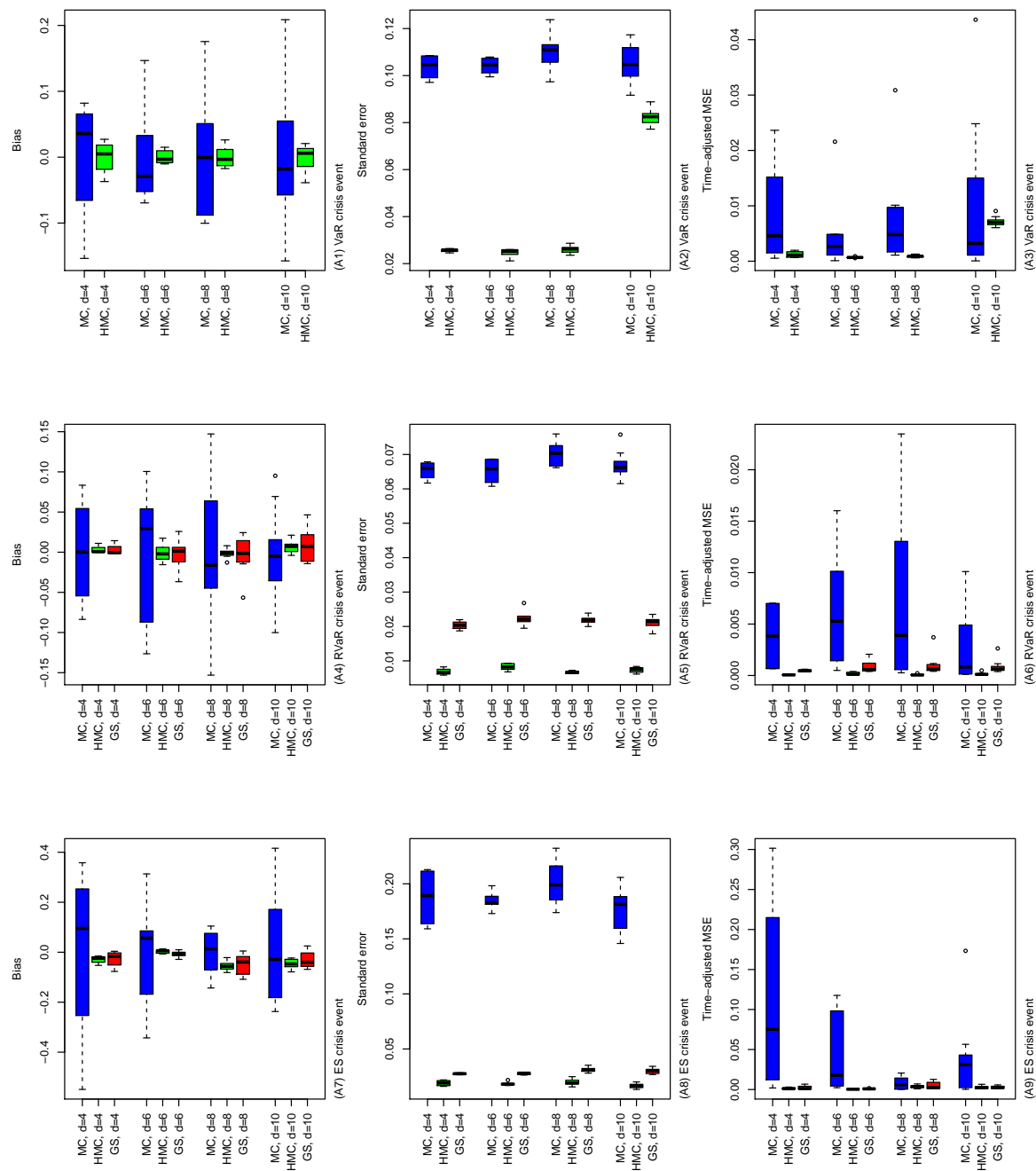
We can then compare the MC and MCMC estimators in terms of bias, standard error, and time-adjusted MSE under the following three scenarios:

- (A) VaR<sub>0.99</sub>, RVaR<sub>0.95,0.99</sub>, and ES<sub>0.99</sub> contributions are estimated by the MC, HMC, and GS methods for dimensions  $d \in \{4, 6, 8, 10\}$ . Note that the GS is applied only to RVaR and ES contributions, not to VaR contributions (same in the other scenarios).
- (B) For  $d = 5$ , VaR <sub>$\alpha$</sub> , RVaR <sub>$\alpha_1^{RVaR}, \alpha_2^{RVaR}$</sub>  and ES <sub>$\alpha$</sub>  contributions are estimated by the MC, HMC, and GS methods for confidence levels  $\alpha^{VaR} \in \{0.9, 0.99, 0.999, 0.9999\}$ ,  $(\alpha_1^{RVaR}, \alpha_2^{RVaR}) \in \{(0.9, 0.9999), (0.9, 0.99), (0.99, 0.999), (0.999, 0.9999)\}$  and  $\alpha^{ES} \in \{0.9, 0.99, 0.999, 0.9999\}$ .
- (C) For  $d = 5$ , VaR<sub>0.9</sub>, RVaR<sub>0.9,0.99</sub> and ES<sub>0.9</sub> contributions are estimated by the MC and HMC methods with the parameters  $(\epsilon_{opt}, T_{opt})$  (determined by Algorithm 4) and  $(\epsilon, T) \in \{(10\epsilon_{opt}, 2T_{opt}), (10\epsilon_{opt}, T_{opt}/2), (\epsilon_{opt}/10, 2T_{opt}), (\epsilon_{opt}/10, T_{opt}/2)\}$ .

In the MC method, the modified VaR contribution  $\mathbb{E}[X|C_{\alpha-\delta, \alpha+\delta}^{RVaR}]$  with  $\delta = 0.01$  is computed. Moreover, if the size of the conditional sample for estimating RVaR and ES contributions is less than 100, then the lower confidence level of the crisis event is subtracted by 0.01, so that at least 100 MC presamples are guaranteed. For the sample paths of the MCMC methods, ACR, ACP, and Hamiltonian errors for the HMC methods were inspected and the convergences of the chains were checked, as in Sections 4.1 and 4.2.

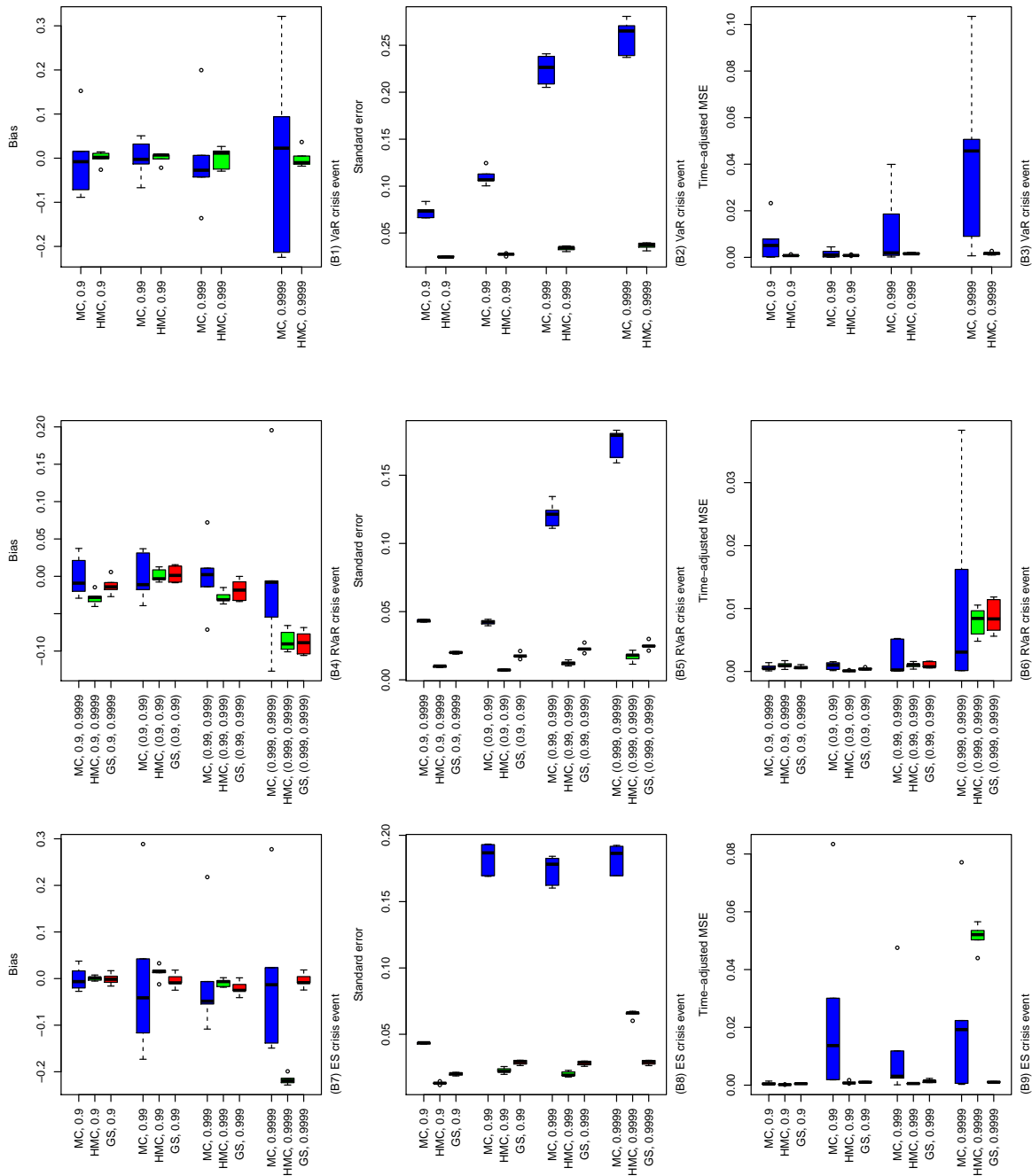
The results of the comparisons of (A), (B), and (C) are summarized in Figures 4–6. In Figure 4, the performance of the MC, HMC, and GS estimators is roughly similar across dimensions from 4 to 10. For all crisis events, the HMC and GS estimators outperform MC in terms of bias, standard error, and time-adjusted MSE. From (A5) and (A8), standard errors of the GS estimators are slightly higher than those of the HMC ones, which result in slightly improved performance of the HMC estimator over the GS in terms of MSE. In Figure 5, bias, standard error, and MSE of the MC estimator tend to increase as the probability of the conditioning set decreases. This is simply because the size of the conditional samples in the MC method decreases proportionally to the probability of the crisis event. On the other hand, the HMC and GS estimators provide a stably better performance than MC since such sample size reduction does not occur. As seen in (B4) to (B9) in the cases of RVaR<sub>0.999,0.9999</sub> and ES<sub>0.9999</sub>, however, if the probability of the conditioning event is too small and/or the distribution of the MC presample is too different from the original conditional distribution of interest, then the parameters of the HMC method determined by Algorithm 4 can be entirely different from the optimal, which leads to a poor performance of the HMC method, as we will see in the next scenario (C). In Figure 6, the HMC method with optimally determined parameters from Algorithm 4 is compared to non-optimal parameter choices. First, the optimal HMC estimator outperforms MC in terms of bias, standard error, and time-adjusted MSE. On the other hand, from the plots in Figure 6, we see that some

of the non-optimal HMC estimators are significantly worse than MC. Therefore, a careful choice of the parameters of the HMC method is required to obtain an improved performance of the HMC method compared to MC.

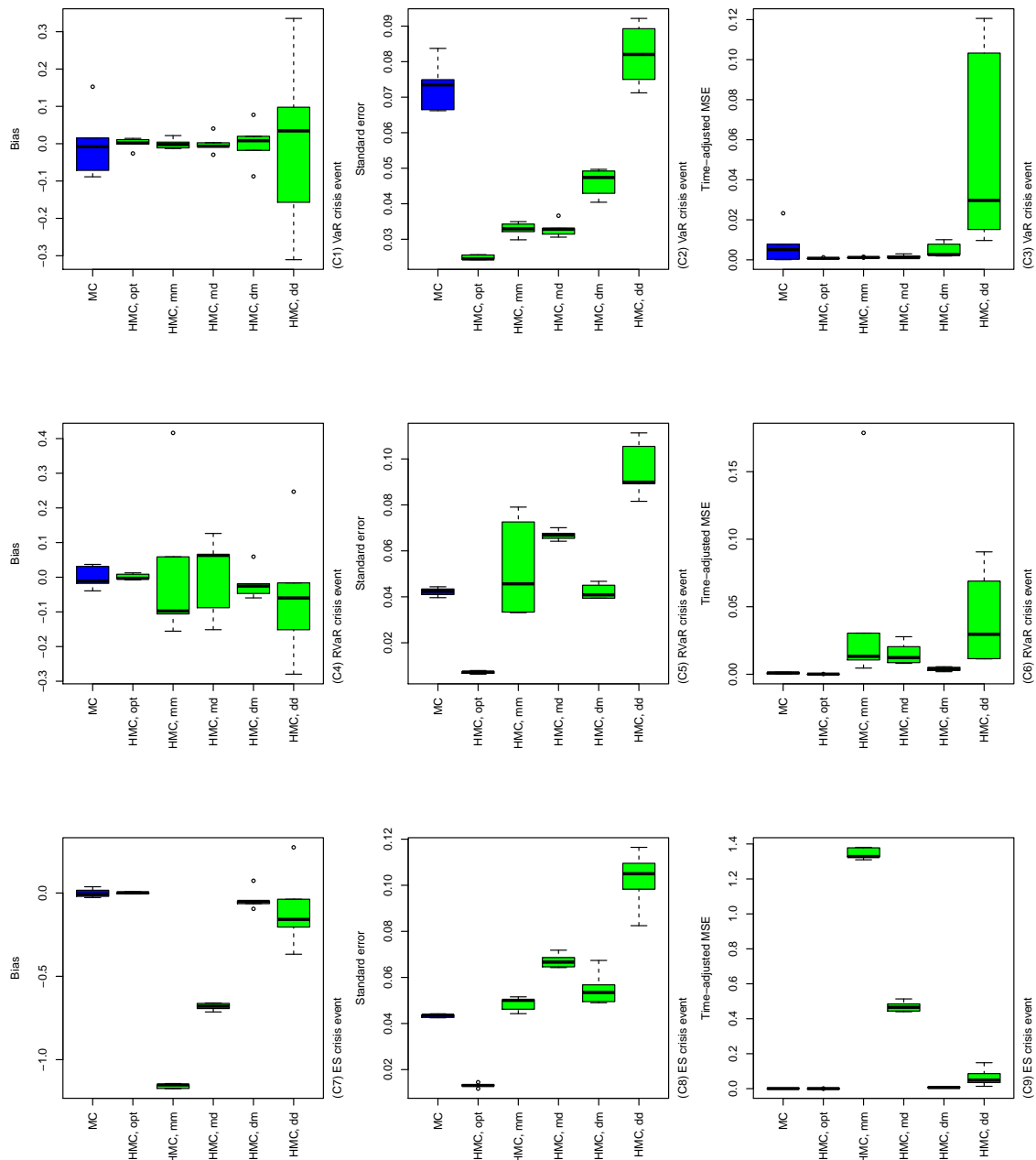


**Figure 4.** Bias (left), standard error (middle) and time-adjusted mean squared error (right) of the MC, HMC, and GS estimators of risk contribution-type systemic risk allocations under VaR<sub>0.99</sub> (top), RVaR<sub>0.95,0.99</sub> (middle), and ES<sub>0.99</sub> (bottom) crisis events. The underlying loss distribution is  $t_\nu(\mu, P)$ , where  $\nu = 6$ ,  $\mu = \mathbf{0}$  and  $P = 1/12 \cdot \mathbf{1}_d \mathbf{1}_d^\top + \text{diag}(11/12)$  for portfolio dimensions  $d \in \{4, 6, 8, 10\}$ . Note that the GS method is applied only to RVaR and ES contributions.





**Figure 5.** Bias (left), standard error (middle), and time-adjusted mean squared error (right) of the MC, HMC, and GS estimators of risk contribution-type systemic risk allocations with the underlying loss distribution  $t_\nu(\mu, P)$ , where  $\nu = 6$ ,  $\mu = \mathbf{0}$ ,  $P = 1/12 \cdot \mathbf{1}_d \mathbf{1}_d^\top + \text{diag}_d(11/12)$  and  $d = 5$ . The crisis event is taken differently, as  $\text{VaR}_{\alpha^{\text{VaR}}}$  (top),  $\text{RVaR}_{\alpha_1^{\text{RVaR}}, \alpha_2^{\text{RVaR}}}$  (middle) and  $\text{ES}_{\alpha^{\text{ES}}}$  (bottom) for confidence levels  $\alpha^{\text{VaR}} \in \{0.9, 0.99, 0.999, 0.9999\}$ ,  $(\alpha_1^{\text{RVaR}}, \alpha_2^{\text{RVaR}}) \in \{(0.9, 0.9999), (0.9, 0.99), (0.99, 0.999), (0.999, 0.9999)\}$ , and  $\alpha^{\text{ES}} \in \{0.9, 0.99, 0.999, 0.9999\}$ . Note that the GS method is applied only to RVaR and ES contributions.



**Figure 6.** Bias (left), standard error (middle), and time-adjusted mean squared error (right) of the MC and HMC estimators of risk contribution-type systemic risk allocations under  $\text{VaR}_{0.9}$ ,  $\text{R VaR}_{0.9,0.99}$ , and  $\text{ES}_{0.9}$  crisis events. The underlying loss distribution is  $t_\nu(\mu, P)$ , where  $\nu = 6$ ,  $\mu = \mathbf{0}$ ,  $P = 1/12 \cdot \mathbf{1}_d \mathbf{1}_d^\top + \text{diag}_d(11/12)$  and  $d = 5$ . The parameters of the HMC method are taken as  $(\epsilon_{\text{opt}}, \epsilon_{\text{opt}})$  determined by Algorithm 4 and  $(\epsilon, T) \in \{(10\epsilon_{\text{opt}}, 2T_{\text{opt}}), (10\epsilon_{\text{opt}}, T_{\text{opt}}/2), (\epsilon_{\text{opt}}/10, 2T_{\text{opt}}), (\epsilon_{\text{opt}}/10, T_{\text{opt}}/2)\}$ . In the labels of the x-axes, each of the five cases  $(\epsilon_{\text{opt}}, \epsilon_{\text{opt}})$ ,  $(10\epsilon_{\text{opt}}, 2T_{\text{opt}})$ ,  $(10\epsilon_{\text{opt}}, T_{\text{opt}}/2)$ ,  $(\epsilon_{\text{opt}}/10, 2T_{\text{opt}})$  and  $(\epsilon_{\text{opt}}/10, T_{\text{opt}}/2)$  is denoted by HMC.opt, HMC.mm, HMC.md, HMC.dm, and HMC.dd, respectively.

### 5. Conclusion, Limitations and Future Work

Efficient calculation of systemic risk allocations is a challenging task, especially when the crisis event has a small probability. To solve this problem for models where a joint loss density is available, we proposed MCMC estimators where a Markov chain is constructed with the conditional loss distribution, given the crisis event as the target distribution. By using HMC and GS, efficient simulation methods from the constrained target distribution were obtained and the resulting MCMC estimator was expected to have a smaller standard error compared to that of the MC estimator. Sample efficiency

is significantly improved, since the MCMC estimator is computed from samples generated directly from the conditional distribution of interest. Another advantage of the MCMC method is that its performance is less sensitive to the probability of the crisis event, and thus to the confidence levels of the underlying risk measures. We also proposed a heuristic for determining the parameters of the HMC method based on the MC presamples. Numerical experiments demonstrated that our MCMC estimators are more efficient than MC in terms of bias, standard error, and time-adjusted MSE. Stability of the MCMC estimation with respect to the probability of the crisis event and efficiency of the optimal parameter choice of the HMC method were also investigated in the experiments.

Based on the results in this paper, our MCMC estimators can be recommended when the probability of the crisis event is too small for MC to sufficiently simulate many samples for a statistical analysis and/or when unbiased systemic risk allocations under the VaR crisis event are required. The MCMC methods are likely to perform well when the dimension of the portfolio is less than or around 10, losses are bounded from the left, and the crisis event is of VaR or RVaR type; otherwise, heavy-tailedness and computational time can become challenging. Firstly, a theoretical convergence result of the HMC method is typically not available when the target distribution is unbounded and heavy-tailed, which is the case when the losses are unbounded and/or the crisis event is of ES type; see the case of the ES crisis event in the empirical study in Section 4.2. Secondly, both the HMC and GS methods suffer from high-dimensional target distributions since the algorithms contain parts of steps where the computational cost linearly increases in dimension. We observed that, in this case, although the MCMC estimator typically improves bias and standard error compared to MC, the improvement vanishes in terms of time-adjusted MSE due to the long computational time of the MCMC method. Finally, multimodality of joint loss distributions and/or the target distribution is also an undesirable feature since full conditional distributions and their inverses (which are required to implement the GS) are typically unavailable in the former case, and the latter case prevents the HMC method from efficiently exploring the entire support of the target distribution. Potential remedies for heavy-tailed and/or high-dimensional target distributions are the HMC method with a non-normal kinetic energy distribution and roll-back HMC; see Appendix B for details. Further investigation of HMC methods and faster methods for determining the HMC parameters are left for future work.

**Supplementary Materials:** An R script for reproducing the numerical experiments conducted in this paper is available at <http://www.mdpi.com/2227-9091/8/1/6/>.

**Author Contributions:** Conceptualization, T.K.; methodology, T.K.; formal analysis, T.K.; investigation, T.K. and M.H.; resources, T.K.; data curation, T.K.; writing—original draft preparation, T.K.; writing—review and editing, M.H.; visualization, T.K. and M.H.; supervision, M.H.; project administration, T.K.; funding acquisition, M.H. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by NSERC through Discovery Grant RGPIN-5010-2015.

**Acknowledgments:** We wish to thank to an associate editor and anonymous referees for their careful reading of the manuscript and their insightful comments.

**Conflicts of Interest:** The authors declare no conflict of interest.

## Abbreviations

The following abbreviations are used in this manuscript:

i.i.d.	Independent and identically distributed
pdf	Probability distribution function
cdf	Cumulative distribution function
ecdf	Empirical cdf
GPD	Generalized Pareto distribution
MSE	Mean squared error
LLN	Law of large numbers
CLT	Central limit theorem
VaR	Value-at-Risk

RVaR	Range VaR
ES	Expected shortfall
MES	Marginal expected shortfall
CoVaR	Conditional VaR
CoES	Conditional ES
MC	Monte Carlo
SMC	Sequential Monte Carlo
MCMC	Markov chain Monte Carlo
ACR	Acceptance rate
ACP	Autocorrelation plot
MH	Metropolis–Hastings
GS	Gibbs sampler
MGS	Metropolized Gibbs sampler
DSGS	Deterministic scan GS
RPGS	Random permutation GS
RSGS	Random scan GS
HMC	Hamiltonian Monte Carlo
RBHMC	Roll-back HMC
RMHMC	Riemannian manifold HMC
ALAE	Allocated loss adjustment expense
P&L	Profit and loss

## Appendix A. Hamiltonian Dynamics with Boundary Reflection

In this appendix, we describe details of the HMC method with boundary reflection, as mentioned in Section 3.3.1. Let  $(\mathbf{h}, v)$  be the hyperplane which the trajectory of the Hamiltonian dynamics hit at  $(\mathbf{x}(t), \mathbf{p}(t))$ . At this time,  $(\mathbf{x}(t), \mathbf{p}(t))$  is immediately replaced by  $(\mathbf{x}(t), \mathbf{p}_r(t))$  where  $\mathbf{p}_r(t)$  is the *reflected momentum* defined by

$$\mathbf{p}_r(t) = \mathbf{p}_{\parallel}(t) - \mathbf{p}_{\perp}(t),$$

where  $\mathbf{p}_{\parallel}(t)$  and  $\mathbf{p}_{\perp}(t)$  are such that  $\mathbf{p}(t) = \mathbf{p}_{\parallel}(t) + \mathbf{p}_{\perp}(t)$  and  $\mathbf{p}_{\parallel}(t)$  and  $\mathbf{p}_{\perp}(t)$  are parallel and perpendicular to the hyperplane  $(\mathbf{h}, v)$ , respectively. Afshar and Domke (2015) and Chevallier et al. (2018) showed that the map  $(\mathbf{x}(t), \mathbf{p}(t)) \mapsto (\mathbf{x}(t), \mathbf{p}_r(t))$  preserves the volume and the Hamiltonian, and that this modified HMC method has the stationary distribution  $\pi$ . As long as the initial position  $\mathbf{x}^{(0)}$  belongs to  $\mathcal{C}$ , the trajectory of the HMC method never violates the constraint  $\mathcal{C}$ . The algorithm of this HMC method with reflection is obtained by replacing the **Leapfrog** function call in Step (3) of Algorithm 3 by Algorithm A1. Accordingly, the parameters of the hyperplanes need to be passed as input to Algorithm 3.

In Step (3-1) of Algorithm A1 the time  $t_m$  at which the trajectory hits the boundary  $(\mathbf{h}_m, v_m)$  is computed. If  $0 < t_m < 1$  for some  $m \in \{1, \dots, M\}$ , then the chain hits the boundary during the dynamics with length  $\epsilon$ . At the smallest time  $t_{m^*}$  among such hitting times, the chain reflects from  $(\mathbf{x}^*, \mathbf{p})$  to  $(\mathbf{x}_r^*, \mathbf{p}_r)$  against the corresponding boundary  $(\mathbf{h}_{m^*}, v_{m^*})$  as described in Step (3-2-1) of Algorithm A1. The remaining length of the dynamics is  $(1 - t_{m^*})\epsilon_{\text{temp}}$  and Step (3) is repeated until the remaining length becomes zero. Other techniques of reflecting the dynamics are discussed in Appendix B.1.

**Algorithm A1** Leapfrog method with boundary reflection.

**Input:** Current state  $(\mathbf{x}(0), \mathbf{p}(0))$ , stepsize  $\epsilon > 0$ , gradients  $\nabla U$  and  $\nabla K$ , and constraints  $(\mathbf{h}_m, v_m)$ ,  $m = 1, \dots, M$ .

**Output:** Updated state  $(\mathbf{x}(\epsilon), \mathbf{p}(\epsilon))$ .

(1) Update  $\mathbf{p}(\epsilon/2) = \mathbf{p}(0) + \epsilon/2 \nabla U(\mathbf{x}(0))$ .

(2) Set  $(\mathbf{x}, \mathbf{p}) = (\mathbf{x}(0), \mathbf{p}(\epsilon/2))$ ,  $\epsilon_{\text{temp}} = \epsilon$ .

(3) **while**  $\epsilon_{\text{temp}} > 0$

(3-1) Compute

$$\mathbf{x}^* = \mathbf{x} + \epsilon_{\text{temp}} \nabla K(\mathbf{p}),$$

$$t_m = (v_m - \mathbf{h}_m^\top \mathbf{x}) / (\epsilon \mathbf{h}_m^\top \mathbf{p}), \quad m = 1, \dots, M.$$

(3-2) **if**  $t_m \in [0, 1]$  for any  $m = 1, \dots, M$ ,

(3-2-1) Set

$$m^* = \operatorname{argmin}\{t_m \mid 0 \leq t_m \leq 1, m = 1, \dots, M\},$$

$$\mathbf{x}_r^* = \mathbf{x}^* - 2 \frac{\mathbf{h}_{m^*}^\top \mathbf{x}^* - v_{m^*}}{\mathbf{h}_{m^*}^\top \mathbf{h}_{m^*}} \mathbf{h}_{m^*},$$

$$\mathbf{p}_r = \frac{\mathbf{x}^* - \mathbf{x} - t_{m^*} \epsilon \mathbf{p}}{\epsilon(1 - t_{m^*})}.$$

(3-2-2) Set  $(\mathbf{x}, \mathbf{p}) = (\mathbf{x}_r^*, \mathbf{p}_r)$  and  $\epsilon_{\text{temp}} = (1 - t_{m^*}) \epsilon_{\text{temp}}$ .

**else**

(3-2-3) Set  $(\mathbf{x}, \mathbf{p}) = (\mathbf{x}^*, \mathbf{p})$  and  $\epsilon_{\text{temp}} = 0$ .

**end if**

**end while**

(4) Set  $\mathbf{x}(\epsilon) = \mathbf{x}$  and  $\mathbf{p}(\epsilon) = \mathbf{p} + \frac{\epsilon}{2} \nabla U(\mathbf{x})$ .

**Appendix B. Other MCMC Methods**

In this appendix, we introduce some advanced MCMC techniques potentially applicable to the problem of estimating systemic risk allocations.

*Appendix B.1. Roll-Back HMC*

Yi and Doshi-Velez (2017) proposed *roll-back HMC (RBHMC)*, in which the indicator function  $\mathbf{1}_{\{\mathbf{x} \in \mathcal{C}\}}$  in the target distribution (5) is replaced by a smooth sigmoid function so that the Hamiltonian dynamics naturally move back inwards when the trajectory violates the constraints. HMC with reflection presented in Section 3.3.1 requires to check  $M$  boundary conditions at every iteration of the Hamiltonian dynamics. In our problem the number  $M$  linearly increases with the dimension  $d$  in the case of pure losses, which leads to a linear increase in the computational cost. The RBHMC method avoids such explicit boundary checks, and thus can reduce the computational cost of the HMC method with constrained target distributions. Despite saving computational time, we observed that the RBHMC method requires a careful choice of the stepsize  $\epsilon > 0$  and the smoothness parameter of the sigmoid function involved, and we could not find any guidance on how to choose them to guarantee a stable performance.

*Appendix B.2. Riemannian Manifold HMC*

Livingstone et al. (2019) indicated that non-normal kinetic energy distributions can potentially deal with heavy-tailed target distributions. In fact, the kinetic energy distribution  $F_K$  can even be dependent on the position variable  $\mathbf{x}$ . For example, when  $F_K(\cdot | \mathbf{x}) = \mathbf{N}(\mathbf{0}, G(\mathbf{x}))$  for a positive definite matrix  $G(\mathbf{x}) > 0$  and  $\mathbf{x} \in E$ , the resulting HMC method is known as *Riemannian manifold HMC (RMHMC)* since

this case is equivalent to applying HMC on the Riemannian manifold with metric  $G(\mathbf{x})$ ; see [Girolami and Calderhead \(2011\)](#). Difficulties in implementing RMHMC are in the choice of metric  $G$  and in the simulation of the Hamiltonian dynamics. Due to the complexity of the Hamiltonian dynamics, simple discretization schemes such as the leapfrog method are not applicable, and the trajectory is updated implicitly by solving some system of equations; see [Girolami and Calderhead \(2011\)](#). Various choices of the metric  $G$  are studied in [Betancourt \(2013\)](#), [Lan et al. \(2014\)](#) and [Livingstone and Girolami \(2014\)](#) for different purposes. Simulation of RMHMC is studied, for example, in [Byrne and Girolami \(2013\)](#).

### Appendix B.3. Metropolized Gibbs Samplers

[Müller \(1992\)](#) introduced the *Metropolized Gibbs sampler (MGS)* in which the proposal density  $q$  in the MH kernel is set to be  $q = f_{Y|v_1 \leq h \top Y \leq v_2}$  where  $Y$  has the same marginal distributions as  $X$  but a different copula  $C^q$  for which  $C_{j|-j}^q$  and  $C_{j|-j}^{q,-1}$  are available so that the GS can be applied to simulate this proposal. This method can be used when the inversion method is not feasible since  $C_{j|-j}$  or  $C_{j|-j}^{-1}$  are not available. Following the MH algorithm, the candidate is accepted with the acceptance probability (4), which can be simply written as

$$\alpha(\mathbf{x}, \tilde{\mathbf{x}}) = \min \left\{ \frac{c(\mathbf{F}(\tilde{\mathbf{x}}))c^q(\mathbf{F}(\mathbf{x}))}{c(\mathbf{F}(\mathbf{x}))c^q(\mathbf{F}(\tilde{\mathbf{x}}))}, 1 \right\}.$$

As an example of the MGS, suppose  $C$  is the Gumbel copula, for which the full conditional distributions cannot be inverted analytically. One could then choose the survival Clayton copula as the proposal copula  $C^q$  above. For this choice of copula,  $q_{j|-j}$  is available by the inversion method as discussed in Section 3.4.1. Furthermore, the acceptance probability is expected to be high especially on the upper tail part because the upper threshold copula of  $C$  defined as  $\mathbb{P}(\mathbf{U} > \mathbf{v} | \mathbf{U} > \mathbf{u})$ ,  $\mathbf{v} \in [\mathbf{u}, \mathbf{1}]$ ,  $\mathbf{u} \in [0, 1]^d$ ,  $\mathbf{U} \sim C$  is known to converge to that of a survival Clayton copula when  $\lim u_j \rightarrow \infty$ ,  $j = 1, \dots, d$ ; see [Juri and Wüthrich \(2002\)](#), [Juri and Wüthrich \(2003\)](#), [Charpentier and Segers \(2007\)](#) and [Larsson and Nešlehová \(2011\)](#).

## References

- Acharya, Viral V., Lasse H. Pedersen, Thomas Philippon, and Matthew Richardson. 2017. Measuring systemic risk. *The Review of Financial Studies* 30: 2–47. [\[CrossRef\]](#)
- Adrian, Tobias, and Markus K. Brunnermeier. 2016. Covar. *The American Economic Review* 106: 1705. [\[CrossRef\]](#)
- Afshar, Hadi Mohasel, and Justin Domke. 2015. Reflection, refraction, and hamiltonian monte carlo. In *Advances in Neural Information Processing Systems*. Cambridge: The MIT Press, pp. 3007–15.
- Asimit, Alexandru V., Edward Furman, Qihe Tang, and Raluca Vernic. 2011. Asymptotics for risk capital allocations based on conditional tail expectation. *Insurance: Mathematics and Economics* 49: 310–24. [\[CrossRef\]](#)
- Asimit, Alexandru V., and Jinzhu Li. 2018. Systemic risk: An asymptotic evaluation. *ASTIN Bulletin: The Journal of The IAA* 48: 673–98. [\[CrossRef\]](#)
- Bernardi, Mauro, Fabrizio Durante, and Piotr Jaworski. 2017. Covar of families of copulas. *Statistics & Probability Letters* 120: 8–17.
- Beskos, Alexandros, Natesh Pillai, Gareth Roberts, Jesus-Maria Sanz-Serna, and Andrew Stuart. 2010. The acceptance probability of the hybrid monte carlo method in high-dimensional problems. *AIP Conference Proceedings* 1281: 23–6.
- Beskos, Alexandros, Natesh Pillai, Gareth Roberts, Jesus-Maria Sanz-Serna, and Andrew Stuart. 2013. Optimal tuning of the hybrid monte carlo algorithm. *Bernoulli* 19: 1501–34. [\[CrossRef\]](#)
- Betancourt, Michael. 2012. Cruising the simplex: Hamiltonian monte carlo and the dirichlet distribution. *AIP Conference Proceedings* 1443: 157–64.
- Betancourt, Michael. 2013. A general metric for riemannian manifold hamiltonian monte carlo. In *International Conference on Geometric Science of Information*. New York: Springer, pp. 327–34.
- Betancourt, Michael. 2016. Identifying the optimal integration time in hamiltonian monte carlo. *arXiv*. [\[CrossRef\]](#)
- Betancourt, Michael. 2017. A conceptual introduction to hamiltonian monte carlo. *arXiv*. [\[CrossRef\]](#)

- Betancourt, Michael, Simon Byrne, and Mark Girolami. 2014. Optimizing the integrator step size for hamiltonian monte carlo. *arXiv*. [CrossRef]
- Byrne, Simon, and Mark Girolami. 2013. Geodesic monte carlo on embedded manifolds. *Scandinavian Journal of Statistics* 40: 825–45. [CrossRef] [PubMed]
- Cambou, Mathieu, Marius Hofert, and Christiane Lemieux. 2017. Quasi-random numbers for copula models. *Statistics and Computing* 27: 1307–29. [CrossRef]
- Cances, Eric, Frédéric Legoll, and Gabriel Stoltz. 2007. Theoretical and numerical comparison of some sampling methods for molecular dynamics. *ESAIM: Mathematical Modelling and Numerical Analysis* 41: 351–89. [CrossRef]
- Charpentier, Arthur, and Johan Segers. 2007. Lower tail dependence for archimedean copulas: Characterizations and pitfalls. *Insurance: Mathematics and Economics* 40: 525–32. [CrossRef]
- Chen, Chen, Garud Iyengar, and Ciamac C. Moallemi. 2013. An axiomatic approach to systemic risk. *Management Science* 59: 1373–88. [CrossRef]
- Chevallier, Augustin, Sylvain Pion, and Frédéric Cazals. 2018. Hamiltonian Monte Carlo With Boundary Reflections, and Application To Polytope Volume Calculations. Available online: <https://hal.archives-ouvertes.fr/hal-01919855/> (accessed on 8 January 2020).
- Chib, Siddhartha, and Edward Greenberg. 1995. Understanding the Metropolis–Hastings algorithm. *The American Statistician* 49: 327–35.
- Chiragiev, Arthur, and Zinoviy Landsman. 2007. Multivariate pareto portfolios: Tce-based capital allocation and divided differences. *Scandinavian Actuarial Journal* 2007: 261–80. [CrossRef]
- Christen, J. Andrés, Colin Fox, and Mario Santana-Cibrian. 2017. Optimal direction gibbs sampler for truncated multivariate normal distributions. *Communications in Statistics-Simulation and Computation* 46: 2587–600. [CrossRef]
- Denault, Michel. 2001. Coherent allocation of risk capital. *Journal of Risk* 4: 1–34. [CrossRef]
- Devroye, Luc. 1985. *Non-Uniform Random Variate Generation*. New York: Springer.
- Dhaene, Jan, Luc Henrard, Zinoviy Landsman, Antoine Vandendorpe, and Steven Vanduffel. 2008. Some results on the cte-based capital allocation rule. *Insurance: Mathematics and Economics* 42: 855–63. [CrossRef]
- Dhaene, Jan, Andreas Tsanakas, Emiliano A. Valdez, and Steven Vanduffel. 2012. Optimal capital allocation principles. *Journal of Risk and Insurance* 79: 1–28. [CrossRef]
- Duane, Simon, Anthony D. Kennedy, Brian J. Pendleton, and Duncan Roweth. 1987. Hybrid monte carlo. *Physics Letters B* 195: 216–22. [CrossRef]
- Durmus, Alain, Eric Moulines, and Eero Saksman. 2017. On the convergence of hamiltonian monte carlo. *arXiv* arXiv:1705.00166.
- Fan, Guobin, Yong Zeng, and Woon K. Wong. 2012. Decomposition of portfolio var and expected shortfall based on multivariate copula simulation. *International Journal of Management Science and Engineering Management* 7: 153–60. [CrossRef]
- Frees, Edward W., and Emiliano A. Valdez. 1998. Understanding relationships using copulas. *North American Actuarial Journal* 2: 1–25. [CrossRef]
- Furman, Edward, Alexey Kuznetsov, and Ričardas Zitikis. 2018. Weighted risk capital allocations in the presence of systematic risk. *Insurance: Mathematics and Economics* 79: 75–81. [CrossRef]
- Furman, Edward, and Zinoviy Landsman. 2008. Economic capital allocations for non-negative portfolios of dependent risks. *ASTIN Bulletin: The Journal of the IAA* 38: 601–19. [CrossRef]
- Furman, Edward, Ruodu Wang, and Ričardas Zitikis. 2017. Gini-type measures of risk and variability: Gini shortfall, capital allocations, and heavy-tailed risks. *Journal of Banking & Finance* 83: 70–84.
- Furman, Edward, and Ričardas Zitikis. 2008. Weighted risk capital allocations. *Insurance: Mathematics and Economics* 43: 263–9. [CrossRef]
- Furman, Edward, and Ričardas Zitikis. 2009. Weighted pricing functionals with applications to insurance: An overview. *North American Actuarial Journal* 13: 483–96. [CrossRef]
- Gelfand, Alan E., and Adrian F. M. Smith. 1990. Sampling-based approaches to calculating marginal densities. *Journal of the American Statistical Association* 85: 398–409. [CrossRef]
- Gelfand, Alan E., Adrian F. M. Smith, and Tai-Ming Lee. 1992. Bayesian analysis of constrained parameter and truncated data problems using gibbs sampling. *Journal of the American Statistical Association* 87: 523–32. [CrossRef]
- Geman, Stuart, and Donald Geman. 1984. Stochastic relaxation, gibbs distributions, and the bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 6: 721–41. [CrossRef]

- Geweke, John. 1991. Efficient simulation from the multivariate normal and student-t distributions subject to linear constraints and the evaluation of constraint probabilities. In *Computing Science and Statistics: Proceedings of the 23rd Symposium on the Interface*. Fairfax: Interface Foundation of North America, Inc., pp. 571–8.
- Geyer, Charles. 2011. Introduction to markov chain monte carlo. In *Handbook of Markov Chain Monte Carlo*. New York: Springer, pp. 3–47.
- Girardi, Giulio, and A. Tolga Ergün. 2013. Systemic risk measurement: Multivariate garch estimation of covar. *Journal of Banking & Finance* 37: 3169–80.
- Girolami, Mark, and Ben Calderhead. 2011. Riemann manifold langevin and hamiltonian monte carlo methods. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 73: 123–214. [CrossRef]
- Glasserman, Paul. 2005. Measuring marginal risk contributions in credit portfolios. *Journal of Computational Finance* 9: 1–41. [CrossRef]
- Glasserman, Paul. 2013. *Monte Carlo Methods in Financial Engineering*. New York: Springer.
- Glasserman, Paul, and Jingyi Li. 2005. Importance sampling for portfolio credit risk. *Management Science* 51: 1643–56. [CrossRef]
- Gourieroux, Christian, and Alain Monfort. 2013. Allocating systemic risk in a regulatory perspective. *International Journal of Theoretical and Applied Finance* 16: 1350041. [CrossRef]
- Gudmundsson, Thorbjörn, and Henrik Hult. 2014. Markov chain monte carlo for computing rare-event probabilities for a heavy-tailed random walk. *Journal of Applied Probability* 51: 359–76. [CrossRef]
- Gupta, Sourendu, A. Irbac, Frithjof Karsch, and Bengt Petersson. 1990. The acceptance probability in the hybrid monte carlo method. *Physics Letters B* 242: 437–43. [CrossRef]
- Hastings, W. Keith. 1970. Monte carlo sampling methods using markov chains and their applications. *Biometrika* 57: 97–109. [CrossRef]
- Hofert, Marius, Ivan Kojadinovic, Martin Mächler, and Jun Yan. 2018. *Elements of Copula Modeling with R*. New York: Springer Use R! Series. [CrossRef]
- Hoffman, Matthew D., and Andrew Gelman. 2014. The no-u-turn sampler: Adaptively setting path lengths in hamiltonian monte carlo. *Journal of Machine Learning Research* 15: 1593–623.
- Hoffmann, Hannes, Thilo Meyer-Brandis, and Gregor Svindland. 2016. Risk-consistent conditional systemic risk measures. *Stochastic Processes and Their Applications* 126: 2014–37. [CrossRef]
- Hogg, Robert V., and Stuart A. Klugman. 2009. *Loss Distributions*. Hoboken: John Wiley & Sons, Volume 249.
- Jaworski, Piotr. 2017. On conditional value at risk (covar) for tail-dependent copulas. *Dependence Modeling* 5: 1–19. [CrossRef]
- Johnson, Alicia A. 2009. Geometric Ergodicity of Gibbs Samplers. Available online: <https://conservancy.umn.edu/handle/11299/53661> (accessed on 8 January 2020).
- Jones, Galin L., Murali Haran, Brian S. Caffo, and Ronald Neath. 2006. Fixed-width output analysis for markov chain monte carlo. *Journal of the American Statistical Association* 101: 1537–47. [CrossRef]
- Juri, Alessandro, and Mario V. Wüthrich. 2002. Copula convergence theorems for tail events. *Insurance: Mathematics and Economics* 30: 405–20. [CrossRef]
- Juri, Alessandro, and Mario V. Wüthrich. 2003. Tail dependence from a distributional point of view. *Extremes* 6: 213–46. [CrossRef]
- Kalkbrener, Michael, Hans Lotter, and Ludger Overbeck. 2004. Sensible and efficient capital allocation for credit portfolios. *Risk* 17: S19–S24.
- Klugman, Stuart A., and Rahul Parsa. 1999. Fitting bivariate loss distributions with copulas. *Insurance: Mathematics and Economics* 24: 139–48. [CrossRef]
- Koike, Takaaki, and Mihoko Minami. 2019. Estimation of risk contributions with mcmc. *Quantitative Finance* 19: 1579–97. [CrossRef]
- Kromer, Eduard, Ludger Overbeck, and Konrad Zilch. 2016. Systemic risk measures on general measurable spaces. *Mathematical Methods of Operations Research* 84: 323–57. [CrossRef]
- Laeven, Roger J. A., and Marc J. Goovaerts. 2004. An optimization approach to the dynamic allocation of economic capital. *Insurance: Mathematics and Economics* 35: 299–319. [CrossRef]
- Lan, Shiwei, Jeffrey Streets, and Babak Shahbaba. 2014. Wormhole hamiltonian monte carlo. In *Twenty-Eighth AAAI Conference on Artificial Intelligence*. Available online: <https://www.aaai.org/ocs/index.php/AAAI/AAAI14/paper/viewPaper/8437> (accessed on 8 January 2020).



- Larsson, Martin, and Johanna Nešlehová. 2011. Extremal behavior of archimedean copulas. *Advances in Applied Probability* 43: 195–216. [CrossRef]
- Leimkuhler, Benedict, and Sebastian Reich. 2004. *Simulating Hamiltonian Dynamics*. Cambridge: Cambridge University Press, Volume 14.
- Liu, Jun S., Wing H. Wong, and Augustine Kong. 1995. Covariance structure and convergence rate of the gibbs sampler with various scans. *Journal of the Royal Statistical Society: Series B (Methodological)* 57: 157–69. [CrossRef]
- Livingstone, Samuel, Michael Betancourt, Simon Byrne, and Mark Girolami. 2016. On the geometric ergodicity of hamiltonian monte carlo. *arXiv*. arXiv:1601.08057.
- Livingstone, Samuel, Michael F. Faulkner, and Gareth O. Roberts. 2019. Kinetic energy choice in hamiltonian/hybrid monte carlo. *Biometrika* 106: 303–19. [CrossRef]
- Livingstone, Samuel, and Mark Girolami. 2014. Information-geometric markov chain monte carlo methods using diffusions. *Entropy* 16: 3074–102. [CrossRef]
- Mainik, Georg, and Eric Schaanning. 2014. On dependence consistency of covar and some other systemic risk measures. *Statistics & Risk Modeling* 31: 49–77.
- McNeil, Alexander J., Rüdiger Frey, and Paul Embrechts. 2015. *Quantitative Risk Management: Concepts, Techniques and Tools*. Princeton: Princeton University Press.
- Metropolis, Nicholas, Arianna W. Rosenbluth, Marshall N. Rosenbluth, Augusta H. Teller, and Edward Teller. 1953. Equation of state calculations by fast computing machines. *The Journal of Chemical Physics* 21: 1087–92. [CrossRef]
- Meyn, Sean P., and Richard L. Tweedie. 2012. *Markov Chains and Stochastic Stability*. New York: Springer.
- Müller, Peter. 1992. Alternatives to the Gibbs Sampling Scheme. Available online: <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.48.5613> (accessed on 8 January 2020).
- Neal, Radford M. 2011. Mcmc using hamiltonian dynamics. *Handbook of Markov Chain Monte Carlo* 2: 2.
- Nelsen, Roger B. 2006. *An Introduction to Copulas*. New York: Springer.
- Nummelin, Esa. 2002. Mc's for mcmc'ists. *International Statistical Review* 70: 215–40. [CrossRef]
- Nummelin, Esa. 2004. *General Irreducible Markov Chains and Non-Negative Operators*. Cambridge: Cambridge University Press.
- Pakman, Ari, and Liam Paninski. 2014. Exact hamiltonian monte carlo for truncated multivariate gaussians. *Journal of Computational and Graphical Statistics* 23: 518–42. [CrossRef]
- Rodriguez-Yam, Gabriel, Richard A. Davis, and Louis L. Scharf. 2004. Efficient gibbs sampling of truncated multivariate normal with application to constrained linear regression. Unpublished manuscript.
- Rosenthal, Jeffrey S. 2011. Optimal proposal distributions and adaptive mcmc. In *Handbook of Markov Chain Monte Carlo*. Edited by Steve Brooks, Andrew Gelman, Galin Jones and Xiao-Li Meng. Boca Raton: CRC Press.
- Ruján, Pál. 1997. Playing billiards in version space. *Neural Computation* 9: 99–122. [CrossRef]
- Siller, Thomas. 2013. Measuring marginal risk contributions in credit portfolios. *Quantitative Finance* 13: 1915–23. [CrossRef]
- Targino, Rodrigo S., Gareth W. Peters, and Pavel V. Shevchenko. 2015. Sequential monte carlo samplers for capital allocation under copula-dependent risk models. *Insurance: Mathematics and Economics* 61: 206–26. [CrossRef]
- Tasche, Dirk. 1995. *Risk Contributions and Performance Measurement*; Working Paper; München: Technische Universität München.
- Tasche, Dirk. 2001. Conditional expectation as quantile derivative. *arXiv*. arXiv:math/0104190.
- Tasche, Dirk. 2008. Capital allocation to business units and sub-portfolios: The euler principle. In *Pillar II in the New Basel Accord: The Challenge of Economic Capital*. Edited by Andrea Resti. London: Risk Books, pp. 423–53.
- Tierney, Luke. 1994. Markov chains for exploring posterior distributions. *The Annals of Statistics* 1994: 1701–28. [CrossRef]
- Vats, Dootika, James M. Flegal, and Galin L. Jones. 2015. Multivariate output analysis for markov chain monte carlo. *arXiv*. arXiv:1512.07713.
- Vernic, Raluca. 2006. Multivariate skew-normal distributions with applications in insurance. *Insurance: Mathematics and Economics* 38: 413–26. [CrossRef]
- Yamai, Yasuhiro, and Toshinao Yoshida. 2002. Comparative analyses of expected shortfall and value-at-risk: Their estimation error, decomposition, and optimization. *Monetary and Economic Studies* 20: 87–121.
- Yi, Kexin, and Finale Doshi-Velez. 2017. Roll-back hamiltonian monte carlo. *arXiv*. [CrossRef]

