Tail Approximation for Credit Risk Portfolios With Heavy-Tailed Risk Factors

Krassimir Kostadinov *

Abstract

We consider a portfolio credit risk model in the spirit of CreditMetrics [15]. The multivariate normally distributed underlying risk factors in that model are replaced by more general multivariate elliptical factors with heavy-tailed marginals, introducing tail-dependence. We consider a full-scale version of the model, i.e. we incorporate not only the default risk, but also rating migrations, credit spread volatility and recovery risk.

We derive an upper bound of the portfolio loss distribution, which is particularly accurate at high loss levels. Given the complexity of our model, we obtain this results using a mixture of analytic techniques and Monte Carlo simulation. We conclude with an approximation of VaR and a new method to determine the contributions of the individual credits to the overall portfolio risk.

*GKAAM, Munich University of Technology, D-85747 Garching, Germany, tel:(++49)(+89)28917991, email: krassi@ma.tum.de, http://www.ma.tum.de/stat

The author is supported financially by a scholarship from DFG (Deutsche Forschungsgemeinschaft) through Graduate Program Applied Algorithmic Mathematics

AMS 2000 Mathematics Subject Classification: 62H99, 65C20, 91B28

JEL Classification: G11, G21

Keywords: heavy-tailed risk factors, marginal risk contributions, portfolio credit risk, saddlepoint, stochastic approximation, tail dependence

1 Introduction

We consider a portfolio credit risk model in the spirit of CreditMetrics [15] and investigate the loss distribution over fixed time horizon T. With respect to the marginal losses, we retain and enhance all features of CreditMetrics [15] and we incorporate not only the default risk, but also rating migrations, credit spread volatility and recovery risk. The dependence structure in the portfolio is given by a set of underlying risk factors which we model by a general multivariate elliptical distribution with heavy-tailed marginals, introducing tail-dependence. Recent empirical studies (see Fortin and Kuzmics [7]) suggest that financial assets tend to have their extreme losses jointly, which is in favour of our model compared to the CreditMetrics [15] model based on multivariate normal (tail-independent) risk factors.

We derive an upper bound of the portfolio loss distribution, which is particularly accurate at high loss levels. Given the complexity of our model, we obtain this results using a mixture of analytic techniques and Monte Carlo simulation. We conclude with an approximation of VaR and a new method to determine the contributions of the individual credits to the overall portfolio risk. Below we present the model in detail.

Let $(\Omega, \mathcal{F}, P)$ be a probability space which carries all random objects in this paper. For $m \in \mathbb{N}$ let $X = (X_1, \dots, X_m)$ be a random vector with discrete marginals, all having the same range $\{1, 2, \dots, K\}$. The primary object of interest is the distribution of the random variable (r.v.)

$$L = \sum_{j=1}^{m} e_j L_j, \tag{1.1}$$

where for $j = 1, \dots, m$:

- e_j is a known positive constant;

- L_j is a real-valued r.v., defined on the probability space $(\Omega, \mathcal{F}, P(\cdot \,|\, X_j))$, where $P(\cdot \,|\, X_j)$ denotes the conditional probability measure.

We assume further:

(A) L_j are conditionally independent, given X;

(B) given X_j, L_j is independent of X_s for $s = 1, \dots, m$, $s \neq j$;

(C) $C_j \leq L_j \leq 1$ a.s. for any outcome X_j, $j = 1, \dots, m$, where $C_j < 1$ are real constants.

In the credit risk framework, L models the loss of a portfolio of m credit risks (loans, bonds or credit derivatives) up to a fixed time horizon T. The constant e_j is the exposure in currency units (i.e. EURO) of credit j in the portfolio. The r.v. X_j is the unknown rating (the credit quality) of credit j at the time horizon T. Rating 1 models default and credit quality increases with the increase of X_j. The r.v. L_j is the loss based on one currency unit for credit j. Its distribution depends on the rating X_j.

Most of the credit risk models used in practice fit within (1.1). For instance, when $K = 2$ (default and non-default rating) and $L_j = \mathrm{I}_{\{X_j=1\}}$, L is the loss of a credit portfolio under the so called 'actuarial valuation' (see Gordy [14], Section 1). With the actuarial valuation one takes care only of the default risk, and the uncertainty in the recovery of a credit in the event of default is ignored. An extension to random recovery rates has been considered by various authors, see for example Bluhm et al. [1], Section 1.1.3. A further extension to multiple ratings is necessary for the so called 'mark-to-market' valuation, see Gordy [14], Section 3, or CreditMetrics [15].

The complexity of model (1.1) is in the joint distribution of $X = (X_1, \ldots, X_m)$. Usually the marginals of $(X_1, \ldots, X_m)$ are calibrated to historical default and rating transition data, see Lando and Skodeberg [20] and Cantor and Hamilton [3] for some of the contemporary methods. We will denote these probabilities by $P(X_j = k) = p_{j,k}$ and

$$P(X_j \leq s) = \sum_{k=1}^{s} p_{j,k} = p_j^s,\ \ s = 1, \ldots, K,\ j = 1, \ldots, m.$$

In order to model the dependence structure of $X = (X_1, \ldots, X_m)$ we introduce the random vector $Y = (Y_1, \ldots, Y_m)$ with continuous marginal distributions G_j and a copula C, i.e. the multivariate d.f. of Y is given by

$$G_Y(y_1, \ldots, y_m) = C(G_1(y_1), \ldots, G_m(y_m))\,. \tag{1.2}$$

The r.v. Y_j, $j = 1, \ldots, m$, is an abstract risk factor, usually it is interpreted as (standardized) asset value of the company standing behind credit j in the portfolio.

Following the approach in CreditMetrics [15], we set for $j = 1, \ldots, m$

$$X_j = k \iff G_j^{-1}(p_j^{k-1}) \ < \ Y_j \ \leq \ G_j^{-1}(p_j^k),\ k = 1, \ldots, K, \tag{1.3}$$

where we interpret $G_j^{-1}(p_j^0) = -\infty$ and $G_j^{-1}(p_j^K) = \infty$.

Thus we reduce the calibration of the distribution of $X = (X_1, \ldots, X_m)$ to the calibration of the marginal default and transition probabilities and the copula of $Y = (Y_1, \ldots, Y_m)$ (see Frey and McNeil [8], Proposition 3.3). For some background on copulas we refer to Embrechts et al. [5] and Joe [18].

We assume also that the assets $Y_1, \ldots, Y_m$ follow a factor model:

$$Y_j = \sum_{l=1}^{p} \alpha_{j,l} W Z_l + \sigma_j W \epsilon_j,\ j = 1, \ldots, m, \tag{1.4}$$

where:

- $Z = (Z_1, \ldots, Z_p) \in N_p(0, \Sigma)$ (p-dimensional multivariate normal d.f. with covariance matrix Σ) are the **common factors**. The matrix Σ is usually calibrated to regional or business sector stock index data;

- W is a positive r.v., independent of Z; it represents a **common shock** affecting simultaneously all credits;

- ϵ_j, $j = 1, \ldots, m$, are i.i.d $N(0, 1)$ **specific factors,** independent of W and Z;

- the constants $\alpha_{j,l} \in \mathbb{R}$ and $\sigma_j > 0$, $j = 1, \ldots, m$, $l = 1, \ldots, p$, are normalized so that $\mathrm{Var}\,[Y_j \,|\, W] = W$.

Given (1.4), $Y \in N_m(0, W\Sigma_Y)$ (multivariate normal variance mixture with mixing variable W, or otherwise called *multivariate elliptical distribution*). The matrix Σ_Y is explicitly available (see for instance Embrechts et al. [5], Theorem 5.2). The copula of Y is an *elliptical copula*. The marginal distributions of Y_j are all one-dimensional normal variance mixtures, i.e. $Y_j \stackrel{d}{=} WZ_0$, where W is defined as above and $Z_0 \in N(0, 1)$, $Z_0 \perp W$. For some background on elliptical distributions and elliptical copulas see Fang et al. [6] and Hult and Lindskog [16].

The simplest special case of (1.4) is the one-factor Gaussian model, obtained when $W = 1$ a.s. and $p = 1$. This model has been investigated by various authors, see for instance Bluhm et al. [1], Section 2.5.1. The popular in practice model CreditMetrics [15] can be obtained from (1.4) by setting $W = 1$ a.s.

In this paper we are particularly interested in model (1.4), when W belongs to the class of distributions with regularly varying tail at infinity, i.e. for all $t > 0$

$$\lim_{w \to \infty} \frac{P\,(W > tw)}{P\,(W > w)} = t^{-\alpha} \tag{1.5}$$

for some $\alpha > 0$. As shown in Hult and Lindskog [16], Theorem 4.3, only in this case Y_i and Y_j, $i \neq j$, exhibit tail dependence, i.e.

$$\lim_{p \to 0} \frac{P\left(Y_i < G_i^{-1}(p), Y_j < G_j^{-1}(p)\right)}{p} > 0\,.$$

Note that by (1.3) the probability of joint defaults is given by

$$P\,(X_i = 1, X_j = 1) = P\left(Y_i < G_i^{-1}(p_i^1), Y_j < G_j^{-1}(p_j^1)\right)$$

and, taking into account that usually the default probabilities p_i^1 and p_j^1 are small, the pairwise tail dependence of assets Y_i and Y_j results in an increased likelihood for simultaneous defaults in the credit portfolio, thus having an important impact on the credit loss distribution, in particular on its tail (see Frey and McNeil [8] for some numerical examples).

The most frequently used model including a r.v. W satisfying (1.5) is the *t-model* ($p \geq 1$, $W = \sqrt{\frac{\nu}{S_\nu}}$, where $S_\nu \in \chi^2_\nu$ – chi-square distribution with ν degrees of freedom). Then $Y \in T_m(0, \Sigma_Y, \nu)$ (multivariate t-distribution with ν degrees of freedom). This means that $\alpha = \nu$ in (1.5).

The calibration of the model to the available data poses a particular challenge, see Daul et al. [4] and Frey and McNeil [9] for some of the available approaches. Once the parameters are estimated, the following two problems are of particular interest in risk management:

(1) Determine the portfolio risk by computing the value of a certain risk measure; the most frequently used in practice being Value-at-Risk (VaR)

$$\mathrm{VaR}(\alpha) = \inf\{x \in \mathbb{R} : P(L \geq x) \leq \alpha\} \tag{1.6}$$

for some (typically small) $0 < \alpha < 1$.

(2) Determine the contributions of the individual credits to the overall portfolio risk. As shown in Tasche [23], the only vector field suitable for risk-adjusted performance measurement is the gradient of the underlying risk measure with respect to the exposures $\mathrm{e} = (e_1, \ldots, e_m)$ as defined in (1.1). The larger the j'th component of the gradient is, the more risky is position j (a small increase in the exposure leads to high increase in the risk of the portfolio as measured by the underlying risk measure). In the sequel we will call the components of the gradient **marginal risk contributions**. As shown in Tasche [23], under some regularity conditions, the marginal risk contributions for VaR are

$$\frac{\partial}{\partial e_j}\mathrm{VaR}(\alpha) = E\left[L_j \mid L = \sum_{j=1}^{m} e_j L_j = \mathrm{VaR}(\alpha)\right], \; j = 1, \ldots, m. \tag{1.7}$$

The usual approach to these two problems, as presented in CreditMetrics [15], is by means of Monte Carlo simulation. This procedure has essentially two steps:

(1) simulate the assets $Y = (Y_1, \ldots, Y_m)$ and calculate the ratings $X = (X_1, \ldots, X_m)$ by means of (1.3);

(2) given $X = (X_1, \ldots, X_m)$, simulate the losses $(L_1, \ldots, L_m)$ as in (1.1) and calculate L.

However, this approach has many shortcomings:

(1) It requires a huge quantity of random numbers, particularly for high-dimensional portfolios ($2mn$ random numbers for a set of n simulations).

(2) Simple Monte Carlo simulation of heavy-tailed r.v.s. like $Y_1, \ldots, Y_m$ produces notoriously poor results (see Glasserman et al. [12] and references therein).

(3) The estimation of rare event probabilities like $P(L \geq \mathrm{VaR}(\alpha))$ is based on a very small proportion of the simulated dataset, i.e. approximately α out of n; hence the results are subject to significant errors.

(4) It is impossible to obtain the marginal risk contributions $\frac{\partial}{\partial e_j}\mathrm{VaR}\,(\alpha)$ (see Bluhm et al., Section 5.2.2.).

Our paper is organized as follows. In section 2 we derive an upper bound of the tail of the portfolio loss distribution. As it is not possible to compute this upper bound explicitly, we use a stochastic approximation algorithm and Monte Carlo simulation. Under weak regularity conditions, we prove a.s. convergence of the proposed algorithm. We obtain an approximation of VaR, and, as a by-product, an approximation of the marginal risk contributions.

In section 3 we give some numerical examples which demonstrate that the proposed method provides an accurate approximation of VaR at low computational costs. Also, we compare our approximation of the marginal risk contributions with a method suggested in Overbeck [22].

2 Upper bound approximation

2.1 Basic results for the portfolio loss L

We note first that due to assumption (C) in (1.1), L has bounded support

$$L_{min} \leq \sum_{j=1}^{m} C_j e_j \leq L = \sum_{j=1}^{m} e_j L_j \leq \sum_{j=1}^{m} e_j = L_{max}\,. \tag{2.1}$$

From now on we exclude some degenerate cases and we suppose that for any $x < L_{max}$ we have $P(L > x) > 0$. This implies, for instance, that if the distribution of L is discrete, then $P(L = L_{max}) > 0$.

The mean of L is

$$E\,[L] = E\,[E\,[L\,|\,X]] = \sum_{j=1}^{m} e_j\,E\,[L_j\,|\,X] = \sum_{j=1}^{m}\sum_{k=1}^{K} e_j p_{j,k} E\,[L_j\,|\,X_j = k]\ .$$

In the following proposition we derive an expression for the moment generating function of L.

Proposition 2.1. *Assume model (1.1) with (1.3) and (1.4). Then the moment generating function* $\varphi\,(\theta) = E\,[\exp\,(\theta L)]$ *exists for every* $\theta \in \mathbb{R}$ *and is given by*

$$\varphi\,(\theta) = E\,[\exp\,(H\,(W, Z, \theta))] \tag{2.2}$$

with

$$H\,(W, Z, \theta) = \sum_{j=1}^{m} \log H_j\,(W, Z, \theta)\ , \tag{2.3}$$

where for $j = 1, \ldots, m$,

$$H_j(W, Z, \theta) = E\left[\exp(\theta e_j L_j) \mid W, Z\right] = \sum_{k=1}^{K} g_{j,k}(W, Z)\varphi_{j,k}(e_j\theta). \tag{2.4}$$

Furthermore, for all $j = 1, \ldots, m$, $k = 1, \ldots, K$

$$g_{j,k}(w, z) = \Phi\left(\frac{G_j^{-1}(p_j^k)}{\sigma_i}\frac{1}{w} - \sum_{l=1}^{p}\frac{\alpha_{j,l}}{\sigma_j}z_l\right) - \Phi\left(\frac{G_j^{-1}(p_j^{k-1})}{\sigma_i}\frac{1}{w} - \sum_{l=1}^{p}\frac{\alpha_{j,l}}{\sigma_j}z_l\right),$$

where we interpret the second term as 0 for $k = 1$ and the first term as 1 for $k = K$; and, finally

$$\varphi_{j,k}(\theta) = E\left[\exp(\theta L_j) \mid X_j = k\right].$$

Proof. Due to (2.1) we obtain immediately for $\theta > 0$

$$\varphi(\theta) = E\left[\exp(\theta L)\right] \leq \exp(\theta L_{max}),$$

therefore $\varphi(\theta)$ exists for every $\theta \geq 0$. Also we have by (2.1) for $\theta < 0$

$$\varphi(\theta) = E\left[\exp(\theta L)\right] \leq \exp(\theta L_{min}),$$

therefore $\varphi(\theta)$ exists for every $\theta \in \mathbb{R}$.
By conditioning on X we have

$$\varphi(\theta) = E\left[\exp(\theta L)\right] = E_X\, E\left[\exp(\theta L) \mid X\right].$$

By assumption (A) in (1.1) L_j, $j = 1, \ldots, m$ are independent, given X. Therefore

$$\varphi(\theta) = E_X\left[\prod_{j=1}^{m} E\left[\exp(\theta e_j L_j) \mid X\right]\right].$$

Since by assumption (B) in (1.1) L_j, given X_j, is independent of X_s for any $j = 1, \ldots, m$ and $s = 1, \ldots, m$, $s \neq j$, we get

$$\varphi(\theta) = E_X\left[\prod_{j=1}^{m} E\left[\exp(\theta e_j L_j) \mid X_j\right]\right].$$

Due to (1.4), given W and Z, the r.v.s Y_j, $j = 1, \ldots, m$, are independent (inherited by the independence of ϵ_j). Therefore X_j, $j = 1, \ldots, m$, are conditionally independent by means of (1.3). Hence, by conditioning on W and Z we get

$$\begin{aligned}\varphi(\theta) &= E_{W,Z}\left[E_X\left[\prod_{j=1}^{m} E\left[\exp(\theta e_j L_j) \mid X_j\right] \mid W,\, Z\right]\right] \\ &= E_{W,Z}\left[\prod_{j=1}^{m} E_{X_j}\left[E\left[\exp(\theta e_j L_j) \mid X_j\right] \mid W,\, Z\right]\right].\end{aligned}$$

Given (1.4),

$$P\left(X_j = k \,|\, W = w,\, Z = z\right) \;\; = P\left(G_j^{-1}(p_j^{k-1}) \le Y_j < G_j^{-1}(p_j^{k}) \,|\, W = w,\, Z = z\right)$$

$$= P\left(G_j^{-1}(p_j^{k-1}) \le \sum_{l=1}^{p} \alpha_{j,l} W Z_l + \sigma_j W \epsilon_j < G_j^{-1}(p_j^{k}) \,|\, W = w,\, Z = z\right)$$
$$= P\left(\frac{G_j^{-1}(p_j^{k-1})}{\sigma_j}\frac{1}{w} - \sum_{l=1}^{p} \frac{\alpha_{j,l}}{\sigma_j} z_l \le \epsilon_j < \frac{G_j^{-1}(p_j^{k})}{\sigma_j}\frac{1}{w} - \sum_{l=1}^{p} \alpha_{j,l} z_l\right)$$
$$= g_{j,k}(w, z)\,.$$

Therefore

$$E_{X_j}\left[E\left[\exp\left(\theta e_j L_j\right) \,|\, X_j\right] \,|\, W,\, Z\right] = \sum_{k=1}^{K} g_{j,k}(W, Z) \varphi_{j,k}(e_j \theta)$$

and hence

$$\varphi\left(\theta\right) \;\; = E_{W,Z}\left[\prod_{j=1}^{m} \sum_{k=1}^{K} g_{j,k}(W, Z) \varphi_{j,k}(e_j \theta)\right]\,.$$

By assumption (C) in (1.1) $C_j \le L_j \le 1$ and $\varphi_{j,k}(e_j\theta)$ is finite for every $\theta \in \mathbb{R}$, $j = 1, \ldots, m$, $k = 1, \ldots, K$. Therefore we get the required result. □

2.2 Approximation of the tail of the portfolio loss distribution

In order to obtain an upper bound of the portfolio loss distribution we apply Markov's inequality: for any $\theta \ge 0$ and $x \in [E\left[L\right], L_{max})$

$$P\left(L \ge x\right) \le E\left[\exp\left(\theta\left(L - x\right)\right)\right] = \varphi\left(\theta\right) \exp\left(-\theta x\right) =: F\left(\theta, x\right). \tag{2.5}$$

In the next lemma we summarize some of the important properties of $F\left(\theta, x\right)$.

Lemma 2.2. *Let $x \in \left(E\left[L\right], L_{max}\right)$ be fixed. Then the function $F\left(\theta, x\right)$ satisfies the following properties.*

(1) $F\left(\theta, x\right)$ is analytic in θ;

(2) there exists a unique positive point

$$\widehat{\theta} = \widehat{\theta}\left(x\right) = \arg\min_{\theta} F\left(\theta, x\right)\,, \tag{2.6}$$

which is the unique positive solution of the equation $\frac{\partial}{\partial\theta} F\left(\theta, x\right) = 0$;

(3) $F\left(\theta, x\right)$ is strictly decreasing for $\theta < \widehat{\theta}(x)$ and strictly increasing for $\theta > \widehat{\theta}(x)$;

(4) $F(\theta, x) \to \infty$, $\theta \to \infty$;

(5) $\widehat{\theta}\left(x\right) \in (0, \theta_{max}(x))$, where $\theta_{max}(x) < \infty$ satisfies $F(\theta_{max}(x), x) = 1$;

(6) there exist constants $D(x) \geq C(x) > 0$ such that for all θ, $0 < \theta < \infty$, $\theta \neq \widehat{\theta}$

$$C(x) \leq \frac{\frac{\partial}{\partial\theta} F(\theta, x)}{\theta - \widehat{\theta}} \leq D(x) \,. \tag{2.7}$$

Proof. Properties (1), (2) and (3) are standard (see e.g. Jensen [17], Section 1.2 and references therein).

As $x < L_{max}$, there exists $\epsilon > 0$ such that $x + \epsilon < L_{max}$ and hence $P(L > x + \epsilon) > 0$, therefore

$$\lim_{\theta \to \infty} F(\theta, x) \geq \lim_{\theta \to \infty} P(L > x + \epsilon) \exp(\theta\epsilon) = \infty \,,$$

i.e. we get (4).

By (2) we have $0 < \widehat{\theta}(x)$. Note that $F(0, x) = 1$ for every x, hence, due to (3), $F(\widehat{\theta}(x), x) < 1$. As $F(\theta, x)$ is continuous in θ, by means of (4) there exists a point $0 < \theta_{max}(x) < \infty$ which satisfies $F(\theta_{max}(x), x) = 1$. Because of (3), we get $\widehat{\theta}(x) < \theta_{max}(x)$, i.e. (5).

To prove (6) we use a Taylor expansion of $\frac{\partial}{\partial\theta} F(\theta, x)$ around $\widehat{\theta}$ and the fact that $\frac{\partial}{\partial\theta} F(\widehat{\theta}, x) = 0$; we get for $\epsilon = \epsilon(\theta) \in (\min(\widehat{\theta}, \theta), \max(\widehat{\theta}, \theta))$

$$\frac{\frac{\partial}{\partial\theta} F(\theta, x)}{\theta - \widehat{\theta}} = \frac{\partial^2}{\partial\theta^2} F(\epsilon, x) \,.$$

For any fixed $\theta \in \mathbb{R}$ we have

$$\begin{aligned} \frac{\partial^2}{\partial\theta^2} F(\theta, x) &= \frac{\partial^2}{\partial\theta^2} E\left[\exp\left(\theta\left(L - x\right)\right)\right] \\ &= E\left[\left(L - x\right)^2 \exp\left(\theta\left(L - x\right)\right)\right] > 0 \,. \end{aligned}$$

Also, by means of (2.1),

$$\frac{\partial^2}{\partial\theta^2} F(\theta, x) \leq (\max(|L_{min}|, L_{max}) + |x|) \exp\left(\theta\left(L_{max} - x\right)\right) < \infty \,.$$

As by (1) $\frac{\partial^2}{\partial\theta^2} F(\theta, x)$ is continuous, for ϵ in a compact subset of $\mathbb{R}$ the function $\frac{\partial^2}{\partial\theta^2} F(\epsilon, x)$ achieves a minimum and a maximum at some points $\widehat{\epsilon}_{min}$ $\widehat{\epsilon}_{max}$, which is strictly positive, hence there exist constants $C(x)$ and $D(x)$ such that

$$0 < C(x) \leq \frac{\partial^2}{\partial\theta^2} F(\widehat{\epsilon}_{min}, x) \leq \frac{\frac{\partial}{\partial\theta} F(\theta, x)}{\theta - \widehat{\theta}} \leq \frac{\partial^2}{\partial\theta^2} F(\widehat{\epsilon}_{max}, x) \leq D(x) < \infty \,.$$

□

To derive a best upper bound of $P(L \geq x)$ we calculate the **saddlepoint** $\widehat{\theta}$ as defined in (2.6) and we obtain

$$P(L \geq x) \leq F(\widehat{\theta}, x), \;\; x \in (E[L], L_{max}) \tag{2.8}$$

This classical large deviations technique has been successfully applied by Martin et al. [21] in the case of a one-factor Gaussian model ($p = 1$, $W = 1$ in (1.4)). Unfortunately in our case it is not possible to compute $\widehat{\theta}$ explicitly or by simple numerical methods, since the moment generating function $\varphi(\theta)$ is available only in terms of the $(p+1)$-dimensional integral (2.2). As a remedy we develop a Monte Carlo estimator for the saddlepoint $\widehat{\theta}$ and at the same time we obtain an estimator for the best upper bound $F(\widehat{\theta}, x)$ as in (2.8).

The proposed method is in the framework of stochastic approximation algorithms (see Kushner and Yin [19]). More precisely, we approximate the saddlepoint $\widehat{\theta}$, for fixed $x \in (E[L], L_{max})$, by simulating recursively the r.v.s:

$$\theta_{n+1} = \theta_n - a_n T_n, \ n \in \mathbb{N},$$

where θ_1 is an arbitrary positive number,

$$T_n = \frac{\partial}{\partial \theta} \exp\left(H\left(W^{(n)}, Z^{(n)}, \theta_n \right) - \theta_n x \right) \tag{2.9}$$

and $W^{(n)}$ and $Z^{(n)}$ are i.i.d. copies of W and Z, respectively, $a_{n,\, n\in\mathbb{N}}$ is a sequence of positive constants such that

$$\sum_{n=1}^{\infty} a_n = \infty \tag{2.10}$$

$$\sum_{n=1}^{\infty} a_n^2 = A^2 < \infty \tag{2.11}$$

$$\lim_{n\to\infty} a_n = 0\,, \tag{2.12}$$

and $H(W, Z, \theta)$ is defined in (2.3).

In the next theorem we prove that (a modification of) $\theta_n \stackrel{\text{a.s.}}{\to} \widehat{\theta}$. The modification is taken in order to ensure 'stability' of the algorithm, i.e. to avoid θ_n growing to infinity for some $\omega \in \Omega$, see Kushner and Yin [19], section 5.1 for details. Usually in stochastic approximation algorithms stability is achieved by assuring that each iterate θ_n belongs to some compact set which includes the true value $\widehat{\theta}$. In our case this compact set is given by $[0, \theta_{max}(x)]$, where $\theta_{max}(x)$ is the constant from Lemma 2.2 (4). Since the constant $\theta_{max}(x)$ is not explicitly available, we approximate it by a sequence of r.v.s converging a.s. to it.

Theorem 2.3. *Let $x \in (E[L], L_{max})$ be fixed. For $n \in \mathbb{N}$ let $T_n^{(i)}$, $i \in \mathbb{N}$, be i.i.d copies of the r.v. defined in (2.9). Let $a_{n,\, n\in\mathbb{N}}$ be a sequence of positive constants satisfying (2.10), (2.11) and (2.12). Define*

$$\theta_{n+1} = \min_i \left\{ \theta_{n+1}^{(i)} = \theta_n - a_n T_n^{(i)} \,:\, 0 \le \theta_{n+1}^{(i)} \le K_n \right\}, \tag{2.13}$$

where

$$K_n = \sup_{\theta} \{\theta \, : \, F_n(\theta, x) \le 1\}$$

with $F_n(\theta, x)$ *being the empirical counterpart of* $F(\theta, x)$ *as defined in (2.5)*

$$F_n(\theta, x) = \frac{1}{n} \sum_{i=1}^{n} \exp\left(H\left(W^{(i)}, Z^{(i)}, \theta\right) - \theta x\right). \tag{2.14}$$

Then

$$\theta_n \overset{\text{a.s.}}{\to} \widehat{\theta}, \; n \to \infty.$$

Proof. Step 1: We prove that θ_n is finite a.s. for every $n \in \mathbb{N}$.
First we prove by induction that $\theta_n < \infty$ for any fixed n. We have $\theta_1 < \infty$. Assume that $\theta_n < \infty$. Note that

$$|T_n| \;=\; |\frac{\partial}{\partial \theta} H(W, Z, \theta_n) - x| \exp\left(H(W, Z, \theta_n) - \theta_n x\right) .$$

However, since $\theta_n \ge 0$ by means of (2.13), $\exp\left(H(W, Z, \theta_n)\right) \le \exp(\theta_n L_{max})$ because of (2.1); and also by (2.3)

$$\begin{aligned} |\frac{\partial}{\partial \theta} H(W, Z, \theta_n)| &\le \sum_{j=1}^{m} |\frac{\partial}{\partial \theta} \log H_j(W, Z, \theta)| \\ &= \sum_{j=1}^{m} |\frac{e_j E\left[L_j \exp(\theta_n e_j L_j) \,|\, W, Z\right]}{E\left[\exp(\theta_n e_j L_j) \,|\, W, Z\right]}| \\ &\le \sum_{j=1}^{m} \frac{e_j \max(1, |C_j|) \exp(\theta_n e_j)}{\exp(\theta_n e_j C_j)} \end{aligned}$$

by means of assumption (C) in (1.1). Therefore

$$|T_n| \;\le\; K(\theta_n), \tag{2.15}$$

where $K(\theta_n)$ is finite if θ_n is finite. Therefore, by means of (2.13) and (2.15), θ_{n+1} is also finite for a finite n.

Now assume that for some $\omega \in \Omega$ $\theta_n \to \infty$, $n \to \infty$. By the SLLN we have for any $\theta \in \mathbb{R}$

$$F_n(\theta, x) \overset{\text{a.s.}}{\to} E\left[\exp\left(\theta (L - x)\right)\right] = F(\theta, x) \;.$$

Furthermore, $F(\theta, x)$ is continuous in θ and $F_n(\theta, x)$ is a.s. continuous in θ and, by means of Lemma 2.2 (4), $F(\theta, x) \to \infty$ when $\theta \to \infty$. So have for this ω and for some sufficiently large n that $F_n(\theta_n, x) > 1$. Such a value of θ_n is excluded by (2.13). Hence θ_n is finite a.s.

Step 2: We prove that

$$(\theta_n - \widehat{\theta})^2 \overset{\text{a.s.}}{\to} \gamma, \; n \to \infty,$$

where γ is some r.v. with finite mean.

Denote by $\mathcal{F}_n = \sigma\left(\theta_1, \ldots, \theta_n, W_{(1)}, Z_{(1)}, \ldots, W_{(n-1)}, Z_{(n-1)}\right)$, where $W_{(j)}, Z_{(j)}$ are the realizations for which $\theta_{j+1} = \theta_j - a_j T_j^{(i)}$. Denote $T_j = T_j^{(i)}$. From (2.13) we have $\theta_{n+1} - \widehat{\theta} = \theta_n - \widehat{\theta} - a_n T_n$, hence

$$E\left[(\theta_{n+1} - \widehat{\theta})^2 \,|\, \mathcal{F}_n\right] = (\theta_n - \widehat{\theta})^2 + a_n^2 E\left[T_n^2 \,|\, \mathcal{F}_n\right] - 2a_n E\left[T_n(\theta_n - \widehat{\theta}) \,|\, \mathcal{F}_n\right]. \qquad (2.16)$$

Consider first

$$\begin{aligned} E\left[T_n(\theta_n - \widehat{\theta}) \,|\, \mathcal{F}_n\right] &= (\theta_n - \widehat{\theta}) E\left[T_n \,|\, \mathcal{F}_n\right] \\ &= (\theta_n - \widehat{\theta}) E\left[T_n \,|\, \theta_n\right] \\ &= (\theta_n - \widehat{\theta}) \frac{\partial}{\partial \theta} E\left[\exp\left(H\left(W_{(n)}, Z_{(n)}, \theta_n\right) - \theta_n x\right) \,|\, \theta_n\right] \\ &= (\theta_n - \widehat{\theta}) \frac{\partial}{\partial \theta} F\left(\theta_n, x\right) . \end{aligned}$$

Secondly,

$$\begin{aligned} \sum_{k=1}^{n} a_k^2 E\left[T_k^2 \,|\, \mathcal{F}_k\right] &= \sum_{k=1}^{n} a_k^2 E\left[T_k^2 \,|\, \theta_k\right] \\ &\leq \sum_{k=1}^{\infty} a_k^2 E\left[T_k^2 \,|\, \theta_k\right] \\ &\leq K^2 A^2 , \end{aligned} \qquad (2.17)$$

where A^2 is the limiting constant from (2.11) and $K^2 < \infty$ is a constant, independent of θ_k, such that

$$E\left[T_k^2 \,|\, \theta_k\right] \leq K^2 \qquad (2.18)$$

for any $k \in \mathbb{N}$ (the existence of such a constant follows from Step 1 and (2.15)).

Denote

$$M_{n+1} = (\theta_{n+1} - \widehat{\theta})^2 + K^2 A^2 - \sum_{k=1}^{n} a_k^2 E\left[T_k^2 \,|\, \theta_k\right] .$$

From (2.17) we know that $M_n \geq 0$. On the other hand, using (2.16),

$$\begin{aligned} E\left[M_{n+1} \,|\, \mathcal{F}_n\right] &= E\left[(\theta_{n+1} - \widehat{\theta})^2 \,|\, \mathcal{F}_n\right] + K^2 A^2 - \sum_{k=1}^{n} a_k^2 E\left[E\left[T_k^2 \,|\, \theta_k\right] \,|\, \mathcal{F}_n\right] \\ &= M_n - 2a_n E\left[T_n(\theta_n - \widehat{\theta}) \,|\, \mathcal{F}_n\right] \\ &= M_n - 2a_n(\theta_n - \widehat{\theta}) \frac{\partial}{\partial \theta} F\left(\theta_n, x\right) . \end{aligned}$$

By means of (2.7) $(\theta_n-\widehat{\theta})\frac{\partial}{\partial\theta}F(\theta_n,x)>0$ a.s., hence M_n is a non-negative supermartingale. By Doob's limit theorem

$$M_n \stackrel{\text{a.s.}}{\to} M,\ n\to\infty, \tag{2.19}$$

where M is a r.v. with finite mean. As K^2A^2 is some costant and $\sum_{k=1}^{n} a_k^2 E\left[T_k^2\,|\,\theta_k\right]$ is an increasing, but bounded by means of (2.17) sequence, (2.19) implies that $(\theta_n-\widehat{\theta})^2 \stackrel{\text{a.s.}}{\to} \gamma$, where γ is a r.v. with finite mean.

Step 3. Denote $\eta_n = E\,(\theta_n-\widehat{\theta})^2$. We prove that there exists $n_1\in\mathbb{N}$ such that for any $n>n_1$

$$\eta_{n+1}\le (1-a_nC)^2\,\eta_n + a_n^2K^2\,, \tag{2.20}$$

where C is the constant from (2.7) and K is the constant from (2.18).

Denoting $\widetilde{T}_n = T_n - \frac{\partial}{\partial\theta}F(\theta_n,x)$ we obtain from (2.13)

$$\begin{aligned}
\theta_{n+1}-\widehat{\theta} &= \theta_n-\widehat{\theta}-a_n\widetilde{T}_n - a_n\frac{\frac{\partial}{\partial\theta}F(\theta_n,x)}{\theta_n-\widehat{\theta}}(\theta_n-\widehat{\theta})\\
&= \left(1-a_n\frac{\frac{\partial}{\partial\theta}F(\theta_n,x)}{\theta_n-\widehat{\theta}}\right)(\theta_n-\widehat{\theta})-a_n\widetilde{T}_n\,.
\end{aligned}$$

Raising to second power and integrating we get

$$\begin{aligned}
\eta_{n+1} \;\le\; & E\left[\left(1-a_n\frac{\frac{\partial}{\partial\theta}F(\theta_n,x)}{\theta_n-\widehat{\theta}}\right)^2(\theta_n-\widehat{\theta})^2\right]-2a_nE\left[\left(1-a_n\frac{\frac{\partial}{\partial\theta}F(\theta_n,x)}{\theta_n-\widehat{\theta}}\right)(\theta_n-\widehat{\theta})\widetilde{T}_n\right]\\
&+a_n^2E\widetilde{T}_n^2\,.
\end{aligned}$$

Conditioning on $\mathcal{F}_n$ we have

$$\begin{aligned}
E\left[\left(1-a_n\frac{\frac{\partial}{\partial\theta}F(\theta_n,x)}{\theta_n-\widehat{\theta}}\right)(\theta_n-\widehat{\theta})\widetilde{T}_n\right] &= E\left[E\left[\left(1-a_n\frac{\frac{\partial}{\partial\theta}F(\theta_n,x)}{\theta_n-\widehat{\theta}}\right)(\theta_n-\widehat{\theta})\widetilde{T}_n\,|\,\mathcal{F}_n\right]\right]\\
&= E\left[\left(1-a_n\frac{\frac{\partial}{\partial\theta}F(\theta_n,x)}{\theta_n-\widehat{\theta}}\right)(\theta_n-\widehat{\theta})E\left[\widetilde{T}_n\,|\,\mathcal{F}_n\right]\right]\\
&= E\left[\left(1-a_n\frac{\frac{\partial}{\partial\theta}F(\theta_n,x)}{\theta_n-\widehat{\theta}}\right)(\theta_n-\widehat{\theta})E\left[\widetilde{T}_n\,|\,\theta_n\right]\right]\\
&=0\ ,
\end{aligned}$$

because $E\left[\widetilde{T}_n\,|\,\theta_n\right]=E\left[T_n\,|\,\theta_n\right]-E\left[\frac{\partial}{\partial\theta}F(\theta_n,x)\,|\,\theta_n\right]=0$ a.s..

Due to (2.7) we have

$$1-a_n\frac{\frac{\partial}{\partial\theta}F(\theta_n,x)}{\theta_n-\widehat{\theta}}\le 1-a_nC.$$

Also, since a_n converges to 0 and $\frac{\frac{\partial}{\partial\theta}F(\theta_n,x)}{\theta_n-\widehat{\theta}} \leq D$ as in (2.7), we can always select an index n_1, such that $a_n < \frac{1}{D}$ for every $n > n_1$, and hence $1 - a_n\frac{\frac{\partial}{\partial\theta}F(\theta_n,x)}{\theta_n-\widehat{\theta}} > 0$. We get

$$\eta_{n+1} \leq (1 - a_n C)^2 \eta_n + a_n^2 E\widetilde{T}_n^2 .$$

Since $E\widetilde{T}_n^2 = E\left(T_n - \frac{\partial}{\partial\theta}F(\theta_n, x)\right)^2 = Var[T_n] \leq ET_n^2 \leq K^2$ we obtain (2.20).

Step 4. Finally we prove that

$$\lim_{n\to\infty} \eta_n = 0. \tag{2.21}$$

Since a.s. convergence implies convergence in probability, by means of Step 2 we have

$$\lim_{n\to\infty} \eta_n = E\gamma < \infty.$$

On the other hand, since $\gamma = \lim_{n\to\infty}(\theta_n - \widehat{\theta})^2$ is a non-negative r.v., showing (2.21) is enough to prove the theorem.

We have by (2.20), for any sufficiently large $n > n_1$,

$$\sqrt{\eta_{n+1}} \leq (1 - a_n C)\sqrt{\eta_n} + a_n K .$$

We also have from Step 2 that $\lim_{n\to\infty}\sqrt{\eta_n} = \sqrt{E\gamma}$. Now assume that $\sqrt{E\gamma} > \frac{K}{C}$. This implies for some $\epsilon > 0$ $\sqrt{\eta_n} \geq \frac{K}{C} + \epsilon$, for any $n > n_2 = n_2(\epsilon)$. Let $n_2 > n_1$. We get

$$\begin{aligned}\sqrt{\eta_{n+1}} &\leq \sqrt{\eta_n} - a_n C(\frac{K}{C} + \epsilon) + a_n K \\ &= \sqrt{\eta_n} - a_n C\epsilon .\end{aligned}$$

Using the inequality recursively we get

$$\sum_{j=n_2+1}^{n} \left(\sqrt{\eta_{j+1}} - \sqrt{\eta_j}\right) \leq -C\epsilon \sum_{j=n_2+1}^{n} a_j ,$$

which means

$$\sqrt{\eta_{n+1}} \leq \sqrt{\eta_{n_2}} - C\epsilon \sum_{j=n_2+1}^{n} a_j .$$

However, due to (2.10), there exists some index n_3, such that

$$\frac{\sqrt{\eta_{n_2}}}{C\epsilon} < \sum_{j=n_1+1}^{n_3} a_j ,$$

hence $\sqrt{\eta_{n_3}} < 0$, which is a contradiction. Therefore we obtain $E\gamma \leq \frac{K^2}{C^2}$. Hence there exists $n_4 \in \mathbb{N}$ such that for any $n > n_4$ $\eta_n \leq \frac{K^2}{C^2}$.

Going back to (2.20), we have for sufficiently large $n > \max(n_1, n_4)$

$$\begin{aligned}\eta_{n+1} &\leq (1 - a_n C)^2 \eta_n + a_n^2 K^2 \\ &= \eta_n + a_n^2 C^2 \eta_n - 2 a_n C \eta_n + a_n^2 K^2 \\ &\leq \eta_n + a_n^2 C^2 \frac{K^2}{C^2} - 2 a_n C \eta_n + a_n^2 K^2 \\ &= \eta_n - 2 a_n C \eta_n + 2 a_n^2 K^2\end{aligned}$$

Assume the contrary to the hypothesis, that $\eta_n > \epsilon > 0$ for any n larger than some fixed n_5. Then we get

$$\eta_{n+1} \leq \eta_n - 2 a_n C \epsilon + 2 a_n^2 K^2 .$$

Applying the inequality recursively we obtain

$$\sum_{j=n_5+1}^{n} (\eta_{j+1} - \eta_j) \leq \sum_{j=n_5+1}^{n} 2 a_j^2 K^2 - 2C\epsilon \sum_{j=n_5+1}^{n} a_j$$

and therefore

$$\begin{aligned}\eta_{n+1} &\leq \eta_{n_5} + 2K^2 \sum_{j=n_5+1}^{n} a_j^2 - 2C\epsilon \sum_{j=n_5+1}^{n} a_j \\ &\leq \eta_{n_5} + 2K^2 \sum_{j=1}^{\infty} a_j^2 - 2C\epsilon \sum_{j=n_5+1}^{n} a_j .\end{aligned}$$

However, due to (2.10), there exists some index n_6, such that

$$\frac{\eta_{n_5} + 2K^2 A}{2C\epsilon} < \sum_{j=n_5+1}^{n_6} a_j ,$$

hence $\eta_{n_6+1} < 0$, which is a contradiction. Therefore we obtain the required result. □

Next we derive an approximation of the optimal upper bound $F(\widehat{\theta}, x)$ of $P(L \geq x)$ as defined in (2.8).

Proposition 2.4. *Let $x \in (E[L], L_{max})$ be fixed. For $\theta \in \mathbb{R}$ let $F_n(\theta, x)$ be defined as in (2.14). If $\theta_n \overset{\text{a.s.}}{\to} \widehat{\theta}$, $n \to \infty$, then*

$$F_n(\theta_n, x) \overset{\text{a.s.}}{\to} F(\widehat{\theta}, x), \; n \to \infty \tag{2.22}$$

Proof. By the SLLN we have

$$F_n(\theta, x) \overset{\text{a.s.}}{\to} E\left[\exp(\theta(L - x))\right] = F(\theta, x)$$

for every $\theta \in \mathbb{R}$. As $\exp\left(H(W^{(i)}, Z^{(i)}, \theta) - \theta x\right)$ and $F(\theta, x)$ are continuous in θ and $\theta_n \overset{\text{a.s.}}{\to} \widehat{\theta}$, by the continuous mapping theorem we have

$$F_n(\theta_n, x) \overset{\text{a.s.}}{\to} F(\widehat{\theta}, x), \; n \to \infty.$$

□

Remark 2.5. At each simulation step of (2.13), we check if $F_n(\theta_n, x) < 1$. Therefore, the upper bound approximation of $P(L \geq x)$ is available as by-product from the proposed algorithm.

2.3 Approximation of VaR and the marginal risk contributions

In the following lemma we analyse the optimal upper bound $F(\widehat{\theta}(x), x)$ defined in (2.8), i.e.

$$F(\widehat{\theta}(x), x) = \varphi(\widehat{\theta}(x)) \exp(-\widehat{\theta}(x)x) \;, x \in (E[L], L_{max}) \,,$$

where $\varphi(\theta)$ is the moment generating function of L as in (2.2) and $\widehat{\theta}(x)$ is the saddlepoint defined in (2.6).

Lemma 2.6. *Let $F(\widehat{\theta}(x), x)$, $x \in (E[L], L_{max})$ be the function defined in (2.8). Then*

(1) $\widehat{\theta}(x)$ is continuous and strictly increasing in x;

(2) $F(\widehat{\theta}(x), x)$ is continuous and strictly decreasing in x;

(3) the inverse function

$$\widehat{\text{VaR}}(\alpha) = \arg_x \left\{ F(\widehat{\theta}(x), x) = \alpha \right\} \;, \alpha \in (0,1) \tag{2.23}$$

is a well defined and strictly decreasing function;

(4) for every $\alpha \in (0,1)$

$$\text{VaR}(\alpha) \leq \widehat{\text{VaR}}(\alpha) \;. \tag{2.24}$$

Proof. To prove (1) we take into account that $\frac{\partial}{\partial\theta} \log(\varphi(\widehat{\theta}(x))) = x$ (Lemma 2.2, (2)), so we get

$$\frac{\partial}{\partial x}\frac{\partial}{\partial \theta} \log(\varphi(\widehat{\theta}(x))) = 1 \,,$$

therefore $\frac{\partial}{\partial x}\widehat{\theta}(x) \frac{\partial^2}{\partial\theta^2} \log(\varphi(\widehat{\theta}(x))) = 1$. Hence $\frac{\partial}{\partial x}\widehat{\theta}(x) = (\frac{\partial^2}{\partial\theta^2} \log(\varphi(\widehat{\theta}(x))))^{-1} > 0$ by the strict convexity of $\varphi(\theta)$.

To prove (2) we note that $\log(F(\widehat{\theta}(x), x)) = \log(\varphi(\widehat{\theta}(x))) - \widehat{\theta}(x)\, x$. By differentiation we get

$$\begin{aligned} \frac{\partial}{\partial x} \log(F(\widehat{\theta}(x), x)) &= \frac{\frac{\partial}{\partial\theta}\varphi(\widehat{\theta}(x))}{\varphi(\widehat{\theta}(x))} \frac{\partial}{\partial x}\widehat{\theta}(x) - x\frac{\partial}{\partial x}\widehat{\theta}(x) - \widehat{\theta}(x) \\ &= \frac{\partial}{\partial x}\widehat{\theta}(x) \left(\frac{\frac{\partial}{\partial\theta}\varphi(\widehat{\theta}(x))}{\varphi(\widehat{\theta}(x))} - x \right) - \widehat{\theta}(x) \\ &= -\widehat{\theta}(x) < 0 \end{aligned}$$

for any $x \in (E[L], L_{max})$, hence $F(\widehat{\theta}(x), x)$ is strictly decreasing.

Property (2) implies also the existence and the strict monotonicity of the inverse as in (2.23), i.e. (3).

To prove (4) assume the contrary i.e. that $\mathrm{VaR}(\alpha) > \widehat{\mathrm{VaR}}(\alpha)$. Then we have

$$\begin{aligned} P(L \geq \mathrm{VaR}(\alpha)) &\leq P(L \geq \widehat{\mathrm{VaR}}(\alpha)) \\ &\leq F(\widehat{\theta}(\widehat{\mathrm{VaR}}(\alpha)), \widehat{\mathrm{VaR}}(\alpha)) = \alpha \,, \end{aligned}$$

which is a contradiction to the definition of $\mathrm{VaR}(\alpha)$. □

Using the algorithm described in section 2.2, we compute the optimal upper bound $F(\widehat{\theta}(x), x)$ for a number of points $x \in (E[L], L_{max})$ and we find the upper bound approximation of $\mathrm{VaR}(\alpha)$, denoted by $\widehat{\mathrm{VaR}}(\alpha)$ as in (2.23).

We consider also the marginal risk contributions with respect to $\widehat{\mathrm{VaR}}(\alpha)$. We fix α and we define marginal risk contributions as

$$\widehat{A_j} = \frac{\partial}{\partial e_j} \widehat{\mathrm{VaR}}(\alpha; \mathrm{e}) \;\; , \; j = 1, \ldots, m \,, \tag{2.25}$$

where $\mathrm{e} = (e_1, \ldots, e_m) \in \mathbb{R}^m_+$ are the exposures of the individual credits as in (1.1). Note that, for fixed $\alpha \in (0,1)$, $\widehat{\mathrm{VaR}}(\alpha; \mathrm{e})$ is well defined for every $\mathrm{e} \in \mathbb{R}^m_+$. This can be seen from the following arguments.

Since the distribution of L in (1.1) depends on e, its moment generating function depends on e, i.e. we have $\varphi(\theta) = \varphi(\theta; \mathrm{e})$. Due to (2.2) and (2.3) we have also that $\varphi(\theta; \mathrm{e}) = E[\exp(H(W, Z, \theta, \mathrm{e}))]$ is well defined and

$$H(W, Z, \theta, \mathrm{e}) = \sum_{j=1}^{m} \log H_j(W, Z, \theta, e_j) \tag{2.26}$$

with H_j, $j = 1, \ldots, m$ defined in (2.4). Hence, for $x \in (E[L], L_{max})$ we have also that the saddlepoint $\widehat{\theta}(x)$ defined in (2.6) is a function of e, i.e $\widehat{\theta}(x) = \widehat{\theta}(x; \mathrm{e})$. Hence the upper bound $F\left(\widehat{\theta}(x), x\right) = F\left(\widehat{\theta}(x; \mathrm{e}), x; \mathrm{e}\right)$ as in (2.8) is well defined. Therefore, due to lemma 2.6 (3), $\widehat{\mathrm{VaR}}(\alpha; \mathrm{e})$ is well defined for every $\mathrm{e} \in \mathbb{R}^m_+$.

Proposition 2.7. *With the above notations, for any fixed $\alpha \in (0,1)$, the upper bound marginal risk contributions $(\widehat{A}_1, \ldots \widehat{A}_m)$ defined in (2.25) are given by*

$$\widehat{A_j} = E\left[\frac{\exp\left(H(W, Z, \theta, \mathrm{e}) - \theta \widehat{\mathrm{VaR}}(\alpha; \mathrm{e})\right)}{\alpha e_j} \frac{\partial}{\partial \theta} \log H_j(W, Z, \theta, e_j)\right]_{|\theta = \widehat{\theta}(\widehat{\mathrm{VaR}}(\alpha; \mathrm{e}); \mathrm{e})} \,. \tag{2.27}$$

Proof. By the definition of $\widehat{\mathrm{VaR}}(\alpha;\mathrm{e})$ as in (2.23) we have

$$F\left(\widehat{\theta}(\widehat{\mathrm{VaR}}\,(\alpha;\mathrm{e})\,;\mathrm{e}),\widehat{\mathrm{VaR}}\,(\alpha;\mathrm{e})\,;\mathrm{e}\right)=\alpha\,, \tag{2.28}$$

hence

$$\frac{\partial}{\partial e_j}F\left(\widehat{\theta}(\widehat{\mathrm{VaR}}\,(\alpha;\mathrm{e})\,;\mathrm{e}),\widehat{\mathrm{VaR}}\,(\alpha;\mathrm{e})\,;\mathrm{e}\right)=0\,.$$

Therefore, setting $\widehat{\theta}=\widehat{\theta}(\widehat{\mathrm{VaR}}\,(\alpha;\mathrm{e})\,;\mathrm{e})$ we have

$$\begin{aligned}0 \;=\;& \frac{\partial}{\partial\theta}F\left(\theta,\widehat{\mathrm{VaR}}\,(\alpha;\mathrm{e})\,;\mathrm{e}\right)_{|\theta=\widehat{\theta}}\frac{\partial}{\partial e_j}\widehat{\theta}(\widehat{\mathrm{VaR}}\,(\alpha;\mathrm{e})\,;\mathrm{e})\\ &+\left[(\frac{\partial}{\partial e_j}\varphi(\theta;\mathrm{e}))\exp(-\theta\widehat{\mathrm{VaR}}(\alpha;\mathrm{e}))\right]_{|\,\theta=\widehat{\theta}}\\ &-\left[\varphi(\theta;\mathrm{e})\theta(\frac{\partial}{\partial e_j}\widehat{\mathrm{VaR}}(\alpha;\mathrm{e}))\exp(-\theta\widehat{\mathrm{VaR}}(\alpha;\mathrm{e}))\right]_{|\,\theta=\widehat{\theta}}\,,\end{aligned}$$

where the last two summands come from the fact that

$$F(\theta,\widehat{\mathrm{VaR}}\,(\alpha;\mathrm{e})\,;\mathrm{e})=\varphi(\theta;\mathrm{e})\exp(-\theta\widehat{\mathrm{VaR}}(\alpha;\mathrm{e}))\,.$$

Since by (2.6) $\widehat{\theta}(x;\mathrm{e})$ is the point at which the minimum of $F(\theta,x;\mathrm{e})$ with respect to θ is achieved, we have $\frac{\partial}{\partial\theta}F(\theta,x;\mathrm{e})_{|\,\theta=\widehat{\theta}(x;\mathrm{e})}=0$ for every $x\in(E[L],L_{max})$. Therefore we have for $j=1,\ldots,m$

$$\left[\frac{\partial}{\partial e_j}\varphi(\theta;\mathrm{e})-\varphi(\theta;\mathrm{e})\theta\frac{\partial}{\partial e_j}\widehat{\mathrm{VaR}}(\alpha;\mathrm{e})\right]_{|\theta=\widehat{\theta}}=0\,. \tag{2.29}$$

As $\varphi(\theta;\mathrm{e})=E\left[\exp\left(H\left(W,Z,\theta,\mathrm{e}\right)\right)\right]$, we get by means of (2.26) and (2.4)

$$\frac{\partial}{\partial e_j}\varphi(\theta;\mathrm{e})=\frac{\theta}{e_j}E\left[\exp\left(H\left(W,Z,\theta,\mathrm{e}\right)\right)\frac{\partial}{\partial\theta}\log H_j(W,Z,\theta,e_j)\right]\,.$$

On the other hand, by (2.28) we have

$$\varphi(\widehat{\theta}(\widehat{\mathrm{VaR}}\,(\alpha;\mathrm{e})\,;\mathrm{e});\mathrm{e})=\alpha\exp\left(\widehat{\theta}(\widehat{\mathrm{VaR}}\,(\alpha;\mathrm{e})\,;\mathrm{e})\widehat{\mathrm{VaR}}\,(\alpha;\mathrm{e})\right)\,.$$

Substituting this in (2.29) and using the fact that $\widehat{\theta}>0$ we obtain the required result. □

Corollary 2.8. *Denote by* $\widehat{E}\,[\cdot]$ *the expectation under the probability measure defined by*

$$d\widehat{P}\,(L<x)=\frac{\exp(\widehat{\theta}(\widehat{\mathrm{VaR}}(\alpha;\mathrm{e});\mathrm{e})x)}{\varphi(\widehat{\theta}(\widehat{\mathrm{VaR}}(\alpha;\mathrm{e}));\mathrm{e})}dP\,(L<x)\;. \tag{2.30}$$

Then

$$\widehat{A}_j=\widehat{E}\,[L_j]\;,\;\;j=1,\ldots,m\,.$$

Proof. By formula (1.2.2) in Jensen [17], $\widehat{E}[L] = \widehat{\mathrm{VaR}}(\alpha; \mathrm{e})$. On the other hand,

$$\widehat{E}[L] = \sum_{j=1}^{m} e_j \widehat{E}[L_j] \,,$$

and therefore $\widehat{E}[L_j] = \frac{\partial}{\partial e_j} \widehat{E}[L] = \widehat{A}_j$. □

Remark 2.9. The SLLN ensures, for $W^{(i)}, Z^{(i)}$, $i = 1, \ldots, n$ being i.i.d copies of W, Z, and $\widehat{\theta} = \widehat{\theta}(\widehat{\mathrm{VaR}}(\alpha; \mathrm{e}); \mathrm{e})$, that

$$\frac{1}{n} \sum_{i=1}^{n} \frac{\exp(H(W^{(i)}, Z^{(i)}, \widehat{\theta}, \mathrm{e}) - \widehat{\theta}\widehat{\mathrm{VaR}}(\alpha; \mathrm{e}))}{\alpha e_j} \frac{\partial}{\partial \theta} \log H_j(W^{(i)}, Z^{(i)}, \widehat{\theta}, e_j) \overset{\text{a.s.}}{\to} \widehat{A}_j, \;\; n \to \infty \,.$$

Therefore we obtain an estimate for the marginal risk contributions as a by-product from the recursion (2.13).

3 Numerical examples

We consider two examples to demonstrate our method.

Example 3.1. The parameters of the considered model (as in (1.1) with (1.3) and (1.4)) are as follows:

- $m = 100$ credits in the portfolio;
- we generated the exposures $\mathrm{e} = (e_1, \ldots, e_m)$ uniformly on the interval $(1, 25)$;
- we took a rating system with $K = 2$ ratings and we generated the default probabilities $P(X_j = 1) = p_{j,1}, \; j = 1, \ldots, m$ uniformly on the interval $[0.001, 0.02]$;
- the marginal loss distributions are given as $L_j = \mathrm{I}_{\{X_j = 1\}}, \; j = 1, \ldots, m$.

For the dependence structure we use in (1.2) the t-copula with $\nu = 4$ degrees of freedom. We use $p = 21$ common factors and $Z \in N_p(0, I)$ in (1.4). The factor loadings are given by $\alpha_{j1} = 0.7, \; j = 1, \ldots, m, \; \alpha_{jl} = 0.3, \; j = 1, \ldots, m, \; l = j \bmod 10 + 1$ and $\alpha_{jl} = 0.3, \; j = 1, \ldots, m, \; l = [j/10] + 11$. Thus, each credit has a loading of 0.7 on the first factor and loadings 0.3 on two of the next. There are no equivalent credits with respect to the dependence structure. The first factor may be thought of as a global factor, the next ten as regional factors and the last ten as industry factors. This multifactor dependence structure is taken from Glasserman [10], where it is described as particularly hard to deal with.

In figure 1 we compare the tail of the portfolio loss distribution, obtained by Monte Carlo simulation as explained in section 1 to the upper bound approximation as in (2.8), obtained by the new method. We observe that the approximation is quite accurate at high loss levels (i.e. 500-800), but degenerates quickly as we move to the mean of the distribution. Furthermore, by the new method it is possible to obtain estimates and confidence bounds at extremely high loss levels (i.e. 700-1000), where the plain Monte Carlo method degenerates. Note that the new method is computationally more efficient as it requires simulation only of the common factors Z and the global shock W, and not of all random components in the model.

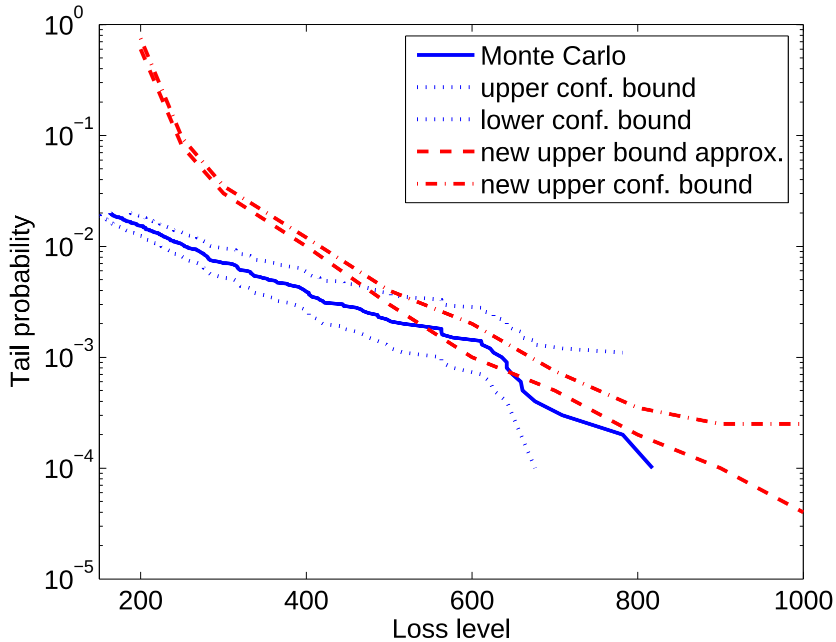

Figure 1: The tail of the portfolio loss distribution obtained by Monte Carlo simulation (10000 simulations) of all random components of L compared to the upper bound approximation obtained by the new method, together with their respective 90% confidence bounds. The approximation is accurate at high loss levels (i.e. 500-800), and stays within the 90% Monte Carlo confidence bounds. At extremely high loss levels (i.e. 700-1000), the new method provides better confidence bounds than the Monte Carlo method.

We further improve the numerical performance of the proposed algorithm by applying importance sampling techniques in the simulation of the common factors Z and the global shock W. We use a classical variance reduction method, namely exponential change of measure, see Glasserman et al. [11]. In the framework of heavy-tailed risk factors, such a technique is not directly applicable, as $E[\exp(\theta W)] = \infty$ for every $\theta > 0$, see Glasserman et al. [12], section 3. Instead, we apply exponential change of measure to the transformed r.v. $S = \frac{1}{W}$. With this method, the accuracy of the approximation at high loss levels can be further improved, as demonstrated in the figure 2.

The second example we consider is simpler, and here we focus not on the absolute portfolio risk, but on the portfolio structure as represented by the marginal risk contributions. We leave the investigation of the impact of the heavy-tailed assumption on the portfolio structure, i.e the comparison between the different copula models, for future work. In

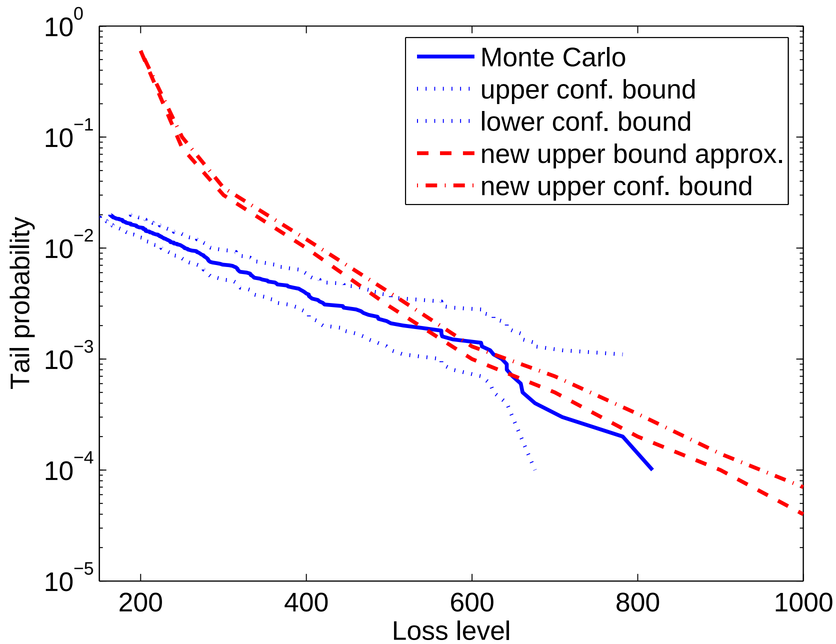

Figure 2: The tail of the portfolio loss distribution obtained by Monte Carlo simulation (10000 simulations) of all random components of L compared to the upper bound approximation obtained by the new method, enhanced by importance sampling, together with their respective 90% confidence bounds. The new method gives accurate information at loss levels where the standard Monte Carlo degenerates (i.e. 700-1000).

this paper we are interested in whether the new method for estimation of marginal risk contributions as in (2.27) provides a good measure for the marginal risks in the portfolio.

We compare our method with an alternative one used for instance in Overbeck [22]. Considering the risk measure Expected Shortfall $\mathrm{ES}(\alpha) = E[L \mid L \geq \mathrm{VaR}(\alpha)]$, one obtains

$$\frac{\partial}{\partial e_j} \mathrm{ES}(\alpha) = E[L_j \mid L \geq \mathrm{VaR}(\alpha)], \ j = 1, \ldots, m, \tag{3.1}$$

see Tasche [23], section 5.3. Let $L^{(i)}$ (and resp. $L_j^{(i)}$, $j = 1, \ldots, m$), $i = 1, \ldots, n$, be i.i.d copies of L (and resp. of L_j, $j = 1, \ldots, m$) as defined in (1.1). Denote by $B_n = \left\{ i = 1, \ldots, n \, : \, L^{(i)} \geq L^{\lceil \alpha n \rceil} \right\}$. Then the SLLN provides estimates for (3.1), i.e.

$$\frac{1}{\# B_n} \sum_{i \in B_n} L_j^{(i)} \stackrel{\text{a.s.}}{\to} \frac{\partial}{\partial e_j} \mathrm{ES}(\alpha), \ n \to \infty .$$

However, as $\# B_n$ increses slowly with the increase of n, extensive Monte Carlo simulation of all random components of L is necessary.

Example 3.2. The parameters of the considered model (as in (1.1) with (1.3) and (1.4)) are as follows:

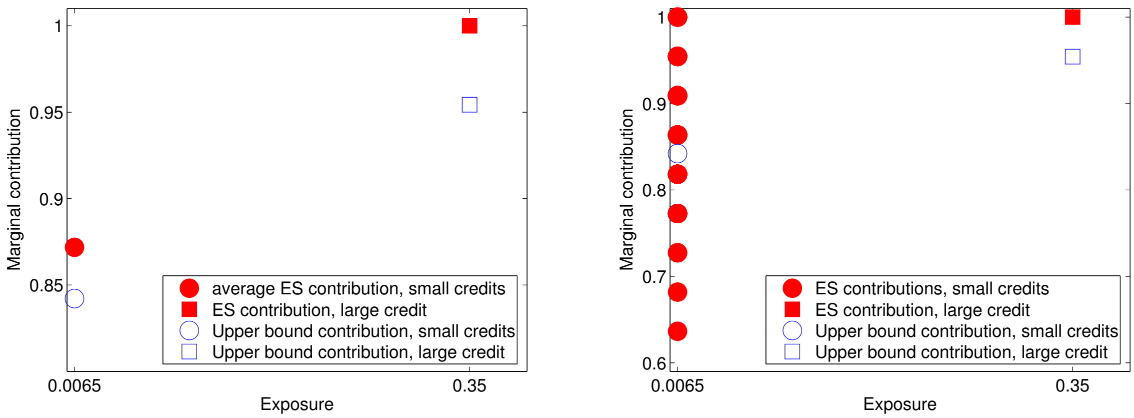

Figure 3: Left figure: the marginal risk contribution of the largest credit and the average marginal risk contribution of the small credits as computed by the ES-method and by the new upper bound method. The results are similar in terms of distance between the small credits and the large one; the overall portfolio risk measured by $\mathrm{ES}(\alpha)$ is higher than the one measured by $\widehat{\mathrm{VaR}}(\alpha)$, as $\widehat{\mathrm{VaR}}(\alpha)$ is an approximation of $\mathrm{VaR}(\alpha) < \mathrm{ES}(\alpha)$.
Right figure: the marginal risk contributions as computed by the two methods. With the ES-method, equivalent credits have different contributions. The new upper bound method avoids this problem.

- $m = 101$ credits in the portfolio;
- we fixed the exposures $e_1 = e_2 = \ldots = e_{m-1} = 0.0065$, however $e_m = 0.35$;
- we took a rating system with $K = 2$ ratings and we fixed the default probabilities $P(X_j = 1) = p_{j,1} = 0.02,\ j = 1, \ldots, m$;
- the marginal loss distributions are given as $L_j = \mathrm{I}_{\{X_j=1\}},\ j = 1, \ldots, m$.

For the dependence structure we use in (1.2) the t-copula with $\nu = 4$ degrees of freedom. We use $p = 1$ common factor with factor loadings $\alpha_{j1} = 0.8,\ j = 1, \ldots, m$ in (1.4).

The parameters in this example are selected in such a way that the portfolio is completely homogeneous, except one of the credits whose exposure is very large. A reasonable method for the computation of the marginal risk contributions should give equal contributions for all credits except the largest, which obviously contributes much more to the the portfolio risk.

We fixed $1 - \alpha = 0.998$ and we computed the risk contributions by the new method (2.27) and by the ES-method (3.1); the results are given on the figure 3. We observe that the two methods provide similar results, however, by the ES-method, due to the error from the Monte Carlo simulation of all random components of L, equivalent credits have different contributions. This problem is avoided by the new method, which uses Monte Carlo simulation only of the common schock W and the common factors Z.

Acknowledgement

I take pleasure in thanking my Ph.D. advisor Claudia Klüppelberg for helpful comments and careful proofreading.

References

[1] Bluhm, C., Overbeck, L. and Wagner, C. (2003): *An Introduction to Credit Risk Modeling*. Chapman & Hall / CRC, Boca Raton.

[2] Bouleau, N. and Lepingle, D. (1994): *Numerical Methods for Stochastic Processes*. John Wiley & Sons, New York.

[3] Cantor, R. and Hamilton, D. (2004): Rating transitions and defaults conditional on watchlist, outlook and rating history. Special comment, Moody's Investment Service. Available at *www.moodys.com*.

[4] Daul, S., De Giorgi, E., Lindskog, F. and McNeil, A. (2003): The grouped t-copula with application to credit risk. *RISK* 16(11), 112-117.

[5] Embrechts, P., Lindskog, F. and McNeil, A. (2001): Modelling dependence with copulas and applications to risk management. Preprint, ETHZ.

[6] Fang, K.T., Kotz, S. and Ng, K.W. (1990): *Symmetric Multivariate and Related Distributions*. Chapman & Hall, London.

[7] Fortin, I. and Kuzmics, C. (2002): Tail dependence in stock return pairs. *International Journal of Intelligent Systems in Accounting, Finance and Management* 11(2), 89-107.

[8] Frey, R. and McNeil, A. (2001): Modelling dependent defaults. Preprint, ETHZ.

[9] Frey, R. and McNeil, A. (2003): Dependent defaults in models of portfolio credit risk. *Journal of Risk* 6(1), 59-92.

[10] Glasserman, P. (2003): Tail approximations for portfolio credit risk. Working paper, Columbia University. Available at *www-1.gsb.columbia.edu/faculty/pgglasserman*.

[11] Glasserman, P., Heidelberger, P. and Shahabuddin, P. (1999): Asymptotically optimal importance sampling for pricing path-dependent options. *Mathematical Finance* 9, 117-152.

[12] Glasserman, P., Heidelberger, P. and Shahabuddin, P. (2000): Portfolio value-at-risk with heavy-tailed risk factors. Working paper, Columbia University. Available at *www-1.gsb.columbia.edu/faculty/pgglasserman*.

[13] Gordy, M. (2000): A comparative anatomy of credit risk models. *Journal of Banking and Finance* 24, 119-149.

[14] Gordy, M. (2001): A risk factor model foundation for rating-based bank capital rules. *Journal of Financial Intermediation* 12, 199-232.

[15] Gupton, G., Finger, C. and Bhatia, M. (1997): *CreditMetrics – Technical document*. J.P.Morgan, New York. Available at *www.riskmetrics.com*.

[16] Hult, H. and Lindskog, F. (2001): Multivariate extremes, aggregation and dependence in elliptical distributions. Research paper, RiskLab. Available at *www.risklab.ch*.

[17] Jensen, J. (1995): *Saddlepoint Approximations*. Oxford University Press, Oxford.

[18] Joe, H. (1997): *Multivariate Models and Dependence Concepts*. Chapman & Hall, New York.

[19] Kushner, H. and Yin, G. (1997): *Stochastic Approximation Algorithms and Applications*. Springer, New York.

[20] Lando, D. and Skodeberg, T. (2000): Analyzing rating transitions and rating drift with continuous observations. *Journal of Banking and Finance* 26, 423-444.

[21] Martin, R., Thompson, K. and Browne, C. (2001): Taking to the saddle. *RISK* 14(6), 91-94.

[22] Overbeck, L. (2000): Allocation of economic capital in loan portfolios. In U. Franke, W. Hardle, G. Stahl, editors, *Measuring Risk in Complex Stochastic Systems*, 62-70, Springer, Berlin.

[23] Tasche, D. (2000): Risk contributions and performance measurement. Working paper, Munich University of Technology. Available at *www-m4.ma.tum.de/m4/Papers*.