A Random Matrix Approach on Credit Risk

Michael C. Münnix,* Rudi Schäfer, and Thomas Guhr
Department of Physics, University of Duisburg-Essen, Duisburg, Germany

We consider a structural model for the estimation of credit risk based on Merton's original model. By using Random Matrix Theory we demonstrate analytically that the presence of correlations severely limits the effect of diversification in a credit portfolio if the correlation are not identical zero. The existence of correlations alters the tails of the loss distribution tremendously, even if their average is zero. Under the assumption of randomly fluctuating correlations, a lower bound for the estimation of the loss distribution is provided.

arXiv:1102.3900v2 [q-fin.RM] 22 Feb 2011

* michael@muennix.com

I. INTRODUCTION

In financial context, the majority of studies in econophysics are dealing with market risk, as many concepts of statistical physics can be directly applied to it. A type of risk that is fundamentally different from the other types of financial risks discussed in the introduction is represented by *credit risk* [1–5]. Modeling credit risk, i.e., the risk that an obligor fails to make a promised payment, is much more complex. As discussed in the previous chapters, the risk of a financial asset, e.g., a stock, is often expressed by its standard deviation. Due to the nearly symmetric shape of the return distribution of, e.g., stocks, by this definition, large positive returns are considered just as risky as large negative returns. An investor who uses volatility as a risk measure acts *risk averse*. However, due to the asymmetric shape of the loss distribution of a credit portfolio, the variance is not suitable as a risk measure in this context.

The primary challenge is to model the loss distribution. It specifies the probability of a certain loss that is caused by obligors that do not pay back their debt. The loss distribution can be described as skewed and leptokurtic [1]. Especially the estimation of its tails, the probability of large losses is of central interest.

The financial crisis of 2008–2009 clearly revealed that an improper estimation of credit risk can lead to drastic effects on the world's economy. The vast underestimation of risks embedded in credits for the subprime housing markets induced a chain reaction that propagated into the worldwide economy. A precise estimation of credit risk is one of the key requirements to avoid a repetition of those events.

As shown by numerical simulations in a previous study of Schäfer et. al. [6], the concept of diversification, i.e. the reduction of risk by increasing the portfolio size, cannot be transferred to credit risk. In a portfolio of credits, such as collateralized debt obligations (CDOs), the effect of diversification is severely limited by the presence of even weak correlations.

In this chapter, we discuss this matter analytically. We estimate the risk embedded in a credit portfolio under the assumption of random correlations with average correlation level zero. This can provide a lower bound for the estimation of the risk in credit portfolios.

II. PHENOMENOLOGY

A credit is a formalized process of lending money. If a *creditor* lends money to an *obligor*, the credit is a written agreement that the obligor will pay back the money. The creditor's reason to lend money is that the obligor does not only have to pay back the amount he borrowed, but also a surplus, typically defined by an interest rate and the expected loss. The obligor takes credits because he wants to raise his capital.

The traditional way of quantifying credit risk is to determine the obligors' credit worthiness. A wide range of rating systems exists that aim at classifying credit worthiness in a system of levels. Prominent examples are the rating agencies *Moody's* and *Standard & Poor*. Rating agencies periodically analyze obligors to provide recent information on their credit worthiness. It is evident, that these rating agencies have a huge impact on the worldwide economy, e.g., if they downgrade the credit rating of a whole country. Criticism arose during the financial crisis of 2008–2009, as the rating principles are not transparent and the agencies are not considered as neutral, i.e., they pursue interests of the country they are based in. Approaches are being made to estimate the *migration risk*, the probability of an obligor being upgraded or downgraded [7–9].

Often, a creditor does not only issue a single credit, but a large number of credits to various institutions and individuals. In this credit portfolio correlations represent an additional observable that influences the risk. For example, during a bad economy, the probability of simultaneous defaults is larger than usual. On the other hand, if the investor arranges his credit portfolio wisely, he might be able to lower the risk of simultaneous defaults, similar to the Markowitz portfolio optimization on stocks [10]. However, this is only feasible if he can quantify the credit portfolio risk properly, which motivates this study.

This paper is organized as follows. In section III, we briefly discuss the different types of credit risk models. A structural credit risk model is considered in section IV, which we demonstrate in an application in section V. We conclude our findings in section VI.

III. TYPES OF CREDIT RISK MODELS

We can distinguish three fundamentally different types of credit risk models (see, e.g., ref. [11]). The models of the first type are referred to as *structural models*. These models have a long history, going back to the work of Black and Scholes [12] and Merton [13]. The *Merton models* assume a zero-coupon debt structure with a fixed time to maturity T. The equity of the company is modeled by a stochastic process, for example describing the price of the company's

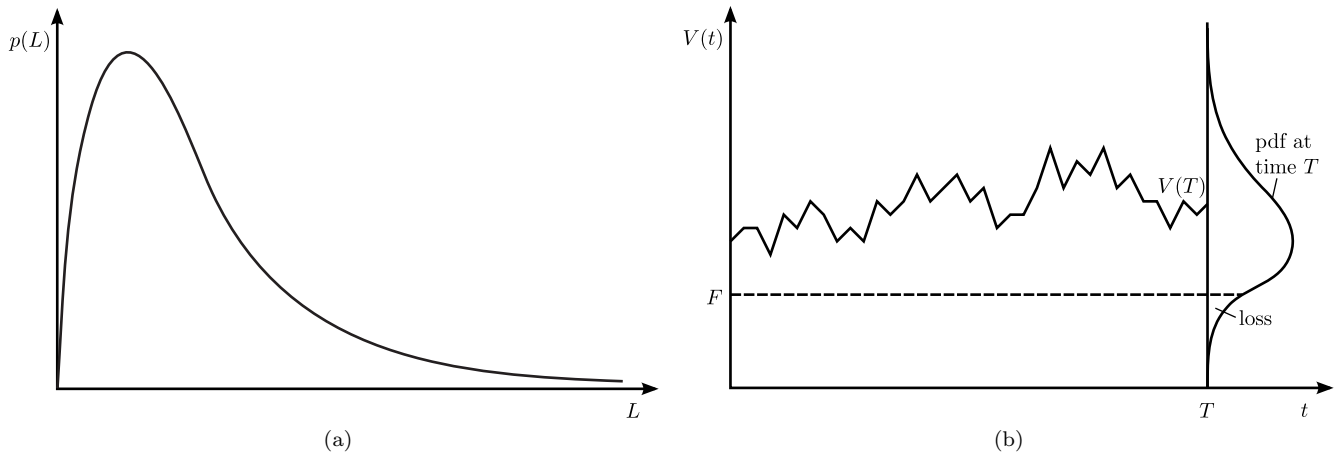


FIG. 1: Sketch of a typical loss distribution $p(L)$. The distribution has a delta peak at $L = 0$ corresponding to the case of non-default. (a) and credit process of the structural model (b).

stock. Thus, it can be seen as a creditor's call option on the obligor's assets. The risk of default and the associated *recovery rate*, the residual payment in case of a loss, are modeled by the company's equity at maturity.

The second type of model are *reduced-form models*. They are frameworks of many common macroscopic observables or risk factors. Despite the large number of input parameters, their functional dependency is usually simple. Default risk and recovery rate are modeled independently. For some well known reduced-form model approaches see refs. [14–18]. These models have to be calibrated with current market data in order to give reliable results. They are widely used in practice throughout the whole financial industry.

First passage models represent the third type of credit risk models. They are a mixed form of the previous two models and were first introduced by Black and Cox [19, 20]. Similar to Merton's model, the market value of a company is modeled through a stochastic process. However, a default occurs whenever the assets go below a certain threshold for the first time. The recovery rates are typically modeled independently, for example, by a reduced-form model [21, 22] or are even assumed to be constant (see, e.g., ref. [11]). Recent approaches aim at improving these models by including the chance of recovery, even if a company's market value is below the threshold [20] and estimating correlations between default probabilities of industry sectors [23].

The latter two models are often implemented in a computer software, for example, *CreditMetrics* by JP Morgan [24], *CreditPortfolioView* by McKinsey & Company [25] or *CreditRisk+* by Credit Suisse [26].

As there can be a strong connection between default risks and recovery rates, the chances of large losses are often underestimated in the reduced-form and first passage models. Structural models do not require this separation.

Because of the high level of abstraction in structural models, they can be characterized as fundamental research in this context. They allow deep insight into the processes of a credit portfolio. However, despite the high level of abstraction, there is a direct application of these models, speculative margin loans: A broker usually grants investors a credit based on his portfolio, a speculative margin loan. For example, if the value of the investor's portfolio value is USD 100 000, the broker might give him a credit of USD 30 000 based on this portfolio. The investor can invest the money on more stocks and try to make more profit than the interest rate on this credit. As the broker maintains his portfolio, he can always take the investor's stocks as a security, if the investor is not willing to pay back his debt. However, if the value of the portfolio falls below USD 30 000 (plus interest), the broker is not able to get back the full amount he lent. A default occurs. Hence, structural models do not only give fundamental insights but also can have high practical relevance. In the next sections, we will develop a structural model with consideration for correlation.

IV. A STRUCTURAL CREDIT RISK MODEL

We develop a structural credit risk model step-by-step. Our model is based on Merton's original model. The aim of our model is to analytically describe the impact of correlations on the losses of a credit portfolio. In our model, a default occurs if the price v_k of the k-th asset is below the face value F_k at maturity time T. The size of the loss then depends on how far $v_k(T)$ is below the face value F_k. The situation is sketched in Fig. 1b. The upright curve at time T represents the probability distribution of the price $v_k(T)$ at time T. The area below F_k in this curve

represents the exposure to default risk. The motivation of the following is to calculate this area and consequently the loss distribution for a portfolio of k assets. We assume that the prices follow a geometric Brownian motion.

A. Average price distribution

For the sake of simplicity, let us first consider the case of a Brownian motion for the average price distribution. This can later easily mapped to the geometric Brownian motion case with a simple substitution. For a geometric Brownian motion, the probability density of the price vector $\mathbf{v}$ of K assets at maturity T is described by

$$p(\mathbf{v}(T),\Sigma) = \left(\frac{1}{2\pi}\right)^K \int \exp\left(-i\omega^\dagger \mathbf{v}(T)\right) \exp\left(-\frac{T}{2}\omega^\dagger \Sigma \omega\right) d[\omega] \tag{1}$$

$$= \left(\frac{1}{\sqrt{2\pi T}}\right)^K \frac{1}{\sqrt{\det(\Sigma)}} \exp\left(-\frac{1}{2T}(\mathbf{v}(T)-\mu T)^\dagger \Sigma^{-1}(\mathbf{v}(T)-\mu T)\right) \ . \tag{2}$$

Here, Σ denotes the covariance matrix and μ is the drift vector.

With Eq. (2) we have the probability distribution of asset prices in the case of a correlated Brownian motion. However, we are not interested in the impact of a specific correlation matrix. Instead we want to estimate the general impact of correlations. To this end, we want to average over all possible correlation matrices and disclose the general statistical behavior of the system. This will enable us to make a profound statistical statement.

We use a random matrix approach to calculate the average price distribution for random correlations where the average correlation level is zero. By averaging over all possible combinations of random variables, we obtain the average price distribution $\langle p(\mathbf{v}(T))\rangle$ under these assumptions.

Thus, we replace the covariance matrix $\mathbf{\Sigma} = \mathbf{A}\mathbf{A}^\dagger$ of the previous section with a random correlation matrix and the matrix of standard derivations $\mathbf{S} = \mathrm{diag}(\sigma_1,\ldots,\sigma_K)$

$$\Sigma_W = SWW^\dagger S \ , \tag{3}$$

where $W \in \mathbb{R}^{K\times N}$ is a random matrix. Hence, the element $(\Sigma_W)_{ij}$ consists of a random number multiplied by $\sigma_i\sigma_j$. The entries of W are Gaussian distributed with

$$p(W) = \left(\sqrt{\frac{N}{2\pi}}\right)^{KN} \exp\left(-\frac{N}{2}\mathrm{tr}\left(WW^\dagger\right)\right) \ . \tag{4}$$

Here, we add the factor N in the nominator because we need the variance in the average price distribution to be independent from N. We will demonstrate that this supplement leads to N-independence later in this section. For $N\to\infty$, we obtain the probability density function of a a unit matrix. This represents the uncorrelated case. For $N = K$, we obtain an invertible covariance matrix with random entries. The case $N < K$ is disregarded as the resulting matrix is not invertible which is usually required for applications in risk management. When inserting this ansatz in Eq. (1), we obtain

$$\langle p(\mathbf{v}(T))\rangle = \int p(W)p(\mathbf{v}(T), SWW^\dagger S)d[W] \tag{5}$$

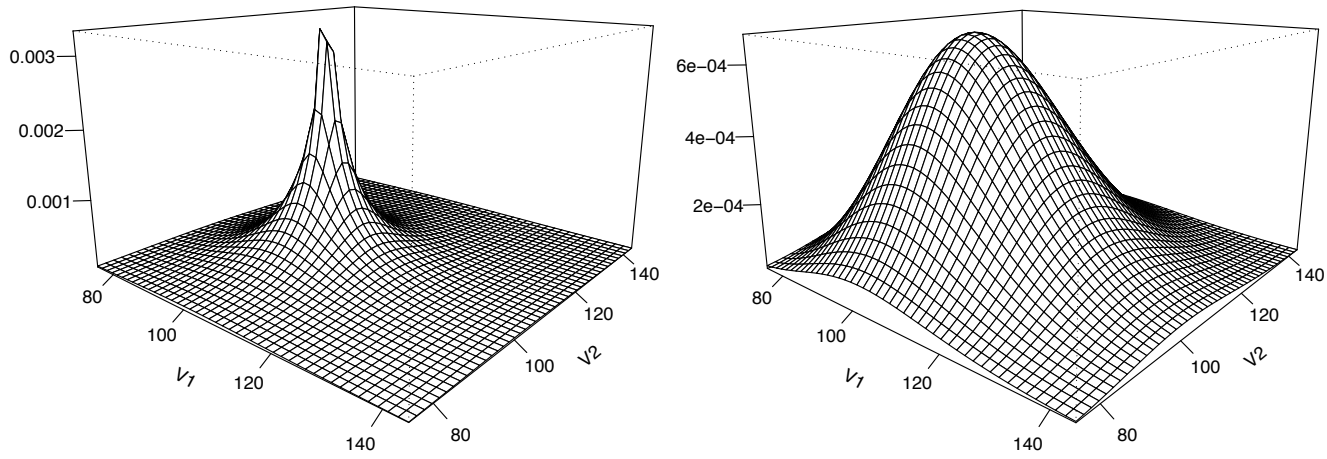
$$= \left(\frac{\sqrt{N}^N}{2\pi}\right)^K \int \exp\left(-i\omega^\dagger \mathbf{v}(T)\right) \frac{1}{\sqrt{\det(NI + TS\omega\omega^\dagger S)}^N} d[\omega] \ , \tag{6}$$

where I denotes the identity matrix. A detailled derivation is given in appendix A. The determinant can also be expressed as

$$\det(NI + TS\omega\omega^\dagger S) = N^K\left(1 + \frac{T}{N}\omega^\dagger SS\omega\right) \tag{7}$$

Hence, we arrive at

$$\langle p(\mathbf{v}(T))\rangle = \left(\frac{1}{2\pi}\right)^K \int \exp\left(-i\omega^\dagger \mathbf{v}(T)\right) \frac{1}{((1+(T/n)\omega^\dagger SS\omega))^{N/2}} d[\omega] \ . \tag{8}$$



(a) N=2 (b) N=100

FIG. 2: Illustration of the average price distribution $\langle p(\mathbf{v}(T))\rangle$ with a geometric Brownian motion for $T = 1$, $K = 2$, $v_{k,0} = 100$, $\mu = 0.05$ and different values for N. All distributions have the identical standard deviation $\hat{\sigma} \approx 16$ ($\sigma = 0.15$). For $N = 2$, we obtain a heavy-tailed distribution while the uncorrelated limit is reached for $N = 100$.

This integral can be solved by identifying the Gamma function [27], as

$$\frac{\Gamma(x)}{a^x} = \int_0^\infty z^{x-1} \exp(-az)\, dz \quad , \quad x > 0,\ a > 0\ . \tag{9}$$

We identify a^{-x} with $((1 + (T/n)\omega^\dagger SS\omega))^{-N/2}$ and obtain

$$\begin{aligned}\langle p(\mathbf{v}(T))\rangle = &\left(\frac{1}{2\pi}\right)^K \frac{1}{\Gamma(N/2)} \left(\prod_{k=1}^K \frac{1}{\sigma_k}\right) \\ &\times \int_0^\infty z^{\left(\frac{N}{2}-1\right)} \exp(-z) \left(\frac{\pi n}{zT}\right)^{\frac{K}{2}} \exp\left(-\frac{N}{4Tz}\sum_{k=1}^K \frac{v_k^2}{\sigma_k^2}\right)\ ,\end{aligned} \tag{10}$$

as worked out appendix B. This integral is a representation of the Bessel function of the second kind $\mathcal{K}$ of the order $(K - N)/2$ [28]. Thus, we obtain the result

$$\begin{aligned}\langle p(\mathbf{v}(T))\rangle = &\left(\sqrt{\frac{N}{2\pi T}}\right)^K \frac{1}{\Gamma(N/2)} \left(\prod_{k=1}^K \frac{1}{\sigma_k}\right) \\ &\times 2^{1-\frac{N}{2}} \left(\sqrt{\frac{N}{T}\sum_{k=1}^K \frac{v_k^2}{\sigma_k^2}}\right)^{\frac{N-K}{2}} \mathcal{K}_{\frac{K-N}{2}}\left(\sqrt{\frac{N}{T}\sum_{k=1}^K \frac{v_k^2}{\sigma_k^2}}\right)\ .\end{aligned} \tag{11}$$

This is the average distribution of $p(\mathbf{v}(T))$ if implying a randomly distributed correlation matrix and an average correlation level of zero. We stated earlier, that we include N in the distribution of the random matrices W in order to render the variance of the average price distribution N-independent. The variances only depend on T and σ_k, as discussed in appendix C.The parameter N is only used to control the correlations. Analogously, in case of a geometric

Brownian motion with a simple substitution,

$$\langle p(\mathbf{v}(T))\rangle = \left(\sqrt{\frac{N}{2\pi T}}\right)^{K} \frac{1}{\Gamma(N/2)} \left(\prod_{k=1}^{K} \frac{1}{\sigma_k v_k}\right) \times 2^{1-\frac{N}{2}} \left(\sqrt{\frac{N}{T}\sum_{k=1}^{K}\frac{\widehat{v}_k^2}{\sigma_k^2}}\right)^{\frac{N-K}{2}} \mathcal{K}_{\frac{K-N}{2}}\left(\sqrt{\frac{N}{T}\sum_{k=1}^{K}\frac{\widehat{v}_k^2}{\sigma_k^2}}\right) \tag{12}$$

with

$$\widehat{v}_k(T) = \ln\left(\frac{v_k(T)}{v_{k,0}}\right) - \left(\mu_k - \frac{\sigma_k^2}{2}\right) T \ . \tag{13}$$

Here, parameter σ_k refers to the standard deviation of the underlying Brownian motion, i.e., the volatility of asset returns. The resulting prices thus have the variance,

$$\hat{\sigma}_k = \sqrt{\exp\left(\sigma_k^2 T + 2\mu\right)\left(\exp\left((\sigma_k^2 T) - 1\right)\right) v_{k,0}^2} \tag{14}$$

Fig. 2 shows the distribution of prices based on a geometric Brownian motion, as given in Eq. (12). The findings are similar to the first case. While we obtain a narrow but heavy-tailed distribution for $N = K$, the distribution becomes more similar to an uncorrelated bivariate log-normal distribution with increasing values of N.

B. Loss distribution

We now turn to the calculation of the loss distribution. In our model, a default occurs if the price at maturity $v_k(T)$, corresponding to the obligor's equity, is below the face value F_k. The size of the loss is given by the difference of F_k and $v_k(T)$. Even if a loss occurs, the creditor might not lose all money that he lent, because the obligor is still able to pay back the amount $v_k(T)$. To compare losses in a portfolio of credits, we have to normalize them by the corresponding face value. We define the loss L_l of the k-th asset as

$$L_l = \begin{cases} \frac{F_k - v_k(T)}{F_k} & v_k(T) < F_k \qquad \text{(default)} \\ 0 & \text{else} \qquad\quad \text{(no default)} \end{cases} \ . \tag{15}$$

When calculating the loss distributions, we have to take into account that for Eq. (15), the prices have to be positive. Thus, for all further considerations, we assume that the underlying process of the price distribution follows a geometric Brownian motion.

When calculating the overall loss of a portfolio, we have to weight each loss by its face value in relation to the sum of all portfolio face values,

$$l = \sum_{k=1}^{K} f_k L_l \quad , \quad f_k = \frac{F_k}{\sum_{l=1}^{K} F_l} \ . \tag{16}$$

We integrate over dimensions of the probability density of prices and filter for those that lead to a given total loss L. By the above stated definitions, we can define a filter for the total loss at maturity time T. In the next step we express the filter using a Fourier transformation. Eventually, we separate those terms that correspond to a default

and those that describe the price above the face value F_k,

$$p(L) = \int_0^\infty d\mathbf{v}(T) p(\mathbf{v}(T)) \delta\left(l - \sum_{k=1}^K f_k L_l\right) \tag{17}$$

$$= \int_0^\infty d\mathbf{v}(T) p(\mathbf{v}(T)) \frac{1}{2\pi} \int_{-\infty}^{+\infty} d\omega \exp\left(-i\omega L + i\omega \sum_{k=1}^K f_k L_l\right) \tag{18}$$

$$= \frac{1}{2\pi} \int_{-\infty}^{+\infty} d\omega \exp\left(-i\omega L\right) \int_0^\infty d\mathbf{v}(T) \exp\left(i\omega \sum_{k=1}^K f_k L_l\right) p(\mathbf{v}(T)) \tag{19}$$

$$= \frac{1}{2\pi} \int_{-\infty}^{+\infty} d\omega \exp\left(-i\omega L\right) \prod_{k=1}^K \left[\int_0^{F_k} dv_k(T) \exp\left(i\omega f_k \left(1 - \frac{v_k}{F_k}\right)\right) + \int_{F_k}^\infty dv_k\right] p(\mathbf{v}(T)) \ . \tag{20}$$

Here, the expression in the squared brackets acts as an operator, because $p(V)$ does not necessarily factorize. We will use this ansatz to calculate the average loss distribution in the next section. However, Eq. (20) can be used to calculate the loss distribution if the actual price distribution is known, i.e., the statistical dependence and the underlying process are estimated. To approach this, it is practical to write Eq. (20) as a combinatorial sum,

$$p(L) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} d\omega \exp\left(-i\omega L\right) \times \sum_{k=1}^K \sum_{j=1}^{\binom{K}{k}} \left[\prod_{l \in \mathrm{Perm}(j,k,K)} \int_0^{F_l} dv_l \exp\left(i\omega f_l \left(1 - \frac{v_l}{F_l}\right)\right) \prod_{\substack{q \in \{1 \dots K\} \\ \setminus \mathrm{Perm}(j,k,K)}} \int_{F_q}^\infty dv_q\right] p(\mathbf{v}(T)) \ , \tag{21}$$

where $\mathrm{Perm}(j,k,K)$ is the j-th permutation of k elements of the set $\{1 \dots K\}$. For example, if $K = 3$ and $k = 2$, we obtain, $\mathrm{Perm}(1,2,3) = \{1,2\}$, $\mathrm{Perm}(2,2,3) = \{2,3\}$ and $\mathrm{Perm}(3,2,3) = \{1,3\}$. However, Eq. (21) might need to be estimated numerically, depending on the complexity of the price distribution $p\mathbf{v}(T))$. In section IV D, we will simplify this combinatorial sum for a homogeneous portfolio and the average price distribution $\langle p(\mathbf{v}(T))\rangle$.

C. Average loss distribution

Now we have developed all necessary tools, to model the average distribution of losses, under the assumption of random correlations and an average correlation level of zero. We start by inserting the average price distribution (B4) into the loss distribution (20),

$$\langle p(L)\rangle = \frac{1}{2\pi\Gamma(N/2)} \int_0^\infty dz \ z^{\left(\frac{N}{2}-1\right)} \exp\left(-z\right) \int_{-\infty}^{+\infty} d\omega \exp\left(-i\omega L\right) r(\omega, z) \ , \tag{22}$$

with

$$r(\omega, z) = \prod_{k=1}^K \left[\int_0^{F_k} dv_k \exp\left(i\omega f_k \left(1 - \frac{v_k}{F_k}\right)\right) + \int_{F_k}^\infty dv_k\right] \times \frac{\sqrt{N}}{2\sigma_k v_k \sqrt{\pi z T}} \exp\left(-\frac{N(\ln(v_k/v_{k,0}) - (\mu - \sigma^2/2)T)^2}{4zT\sigma_k^2}\right) \ . \tag{23}$$

We carry out a second order order approximation of this expression, as carried out in appendix D and arrive at

$$\langle p(L)\rangle = \frac{1}{\sqrt{2\pi}\Gamma(N/2)} \int_0^\infty dz \ z^{\left(\frac{N}{2}-1\right)} \exp\left(-z\right) \frac{1}{\sqrt{\widehat{m}_2(z)}} \exp\left(-\frac{(L - \widehat{m}_1(z))^2}{2\widehat{m}_2(z)}\right) \tag{24}$$

with

$$\widehat{m}_1(z) = \sum_{k=1}^{K} F_k m_{1,k}(z) \tag{25}$$

$$\widehat{m}_2(z) = \sum_{k=1}^{K} F_k^2 (m_{2,k}(z) - m_{1,k}(z)^2) \ . \tag{26}$$

However, the convergence radius of the power series expansion in involved in this approximation is one. Although we consider large portfolios K, i.e., f_k is small, ω runs from $-\infty$ to $+\infty$. This second-order approximation might describe the default terms adequately. However, the non-default terms, corresponding to a delta peak at $L = 0$ require ω to run from $-\infty$ to $+\infty$. Thus, the non-default terms cannot be approximated using this second-order approximation. To address this problem, we will develop an improved approximation in section (IV D).

Due to the complexity of $m_1(z)$ and $m_2(z)$, the dz integral needs to be evaluated numerically. We will present this for an example of a homogeneous portfolio in the next section.

D. Homogenous portfolios

In case of an homogeneous portfolio, in which all credits have the same face value f and the same variance σ^2 the weights can be simplified to

$$F_k = \frac{1}{K} \ . \tag{27}$$

As, $m_{1,k}(z)$ and $m_{1,k}(z)$ become identical for every k, we denote them with $m_1(z)$ and $m_1(z)$ leading to

$$\widehat{m}_1(z) = m_1(z) \tag{28}$$

$$\widehat{m}_2(z) = \frac{1}{K}(m_2(z) - m_1(z)^2) \tag{29}$$

$$m_m(z) = \frac{\sqrt{N}}{2\sigma\sqrt{\pi z T}} \int_0^F \frac{1}{v} \left(1 - \frac{v}{F}\right)^m \exp\left(-\frac{N(\ln(v/v_0) - (\mu - \sigma^2/2)T)^2}{4zT\sigma^2}\right) dv \ . \tag{30}$$

With insertion into Eq. 24, we can calculate the loss distribution for a homogeneous portfolio in the second order approximation. However, the dz integral needs to be evaluated numerically.

1. *Improved approximation*

The second order approximation can be improved by approximating the individual terms of the loss distribution instead of approximating the expression as a whole, similar as discussed in [6]. In case of a homogenous portfolio, the combinatorial sum in Eq. (21) reduces to

$$\begin{aligned}\langle p(L)\rangle = \frac{1}{2\pi\Gamma(N/2)} \int_0^\infty dz \ z^{\left(\frac{N}{2}-1\right)} \exp(-z) \int_{-\infty}^{+\infty} d\omega \exp(-i\omega L) \\ \times \sum_{k=0}^{K} \binom{K}{j} \left(r^{(\mathrm{D})}(\omega, z)\right)^j \left(r^{(\mathrm{ND})}(z)\right)^{K-j} \ ,\end{aligned} \tag{31}$$

with the non-default term

$$\left(r^{(\mathrm{ND})}\right)^{K-j} = \left(\int_F^{\infty} dv \frac{\sqrt{N}}{2\sigma V \sqrt{\pi z T}} \right.$$
$$\left. \times \exp\left(-\frac{N(\ln(v/v_0) - (\mu - \sigma^2/2)T)^2}{4zT\sigma^2} \right) \right)^{K-j} \tag{32}$$
$$= \left(\frac{1}{2} + \frac{1}{2} \mathrm{Erf}\left[\frac{\sqrt{N}(\ln(F/v_0) - (\mu - \sigma^2/2)T)}{2\sigma\sqrt{zT}} \right] \right)^{K-j} \tag{33}$$

and the default term

$$\left(r^{(\mathrm{D})}(\omega, z)\right)^{j} = \left(\int_0^{F} dv \exp\left(\frac{i\omega}{K} \left(1 - \frac{v}{F}\right) \right) \right.$$
$$\left. \times \frac{\sqrt{N}}{2\sigma V \sqrt{\pi z T}} \exp\left(-\frac{N(\ln(v/v_0) - (\mu - \sigma^2/2)T)^2}{4zT\sigma^2} \right) \right)^{j} \tag{34}$$

The approximation follows the same principles as in the previous section, resulting in

$$\int_{-\infty}^{+\infty} d\omega \exp\left(-i\omega L\right) \left(r^{(\mathrm{D})}(\omega, z)\right)^{j}$$
$$= \int_{-\infty}^{+\infty} d\omega \exp\left(i\omega \left(\frac{j}{K} m_1(z) - l \right) - \frac{\omega^2 j}{2K^2} \left(m_2(z) - m_1(z)^2 \right) \right) \tag{35}$$
$$= \sqrt{\frac{2\pi K^2}{j\left(m_2(z) - m_1(z)^2\right)}} \exp\left(-\frac{(LK - j m_1(z))^2}{2j\left(m_2(z) - m_1(z)^2\right)} \right) \ . \tag{36}$$

In this approximation, the non-default terms given by Eq. (33), can be calculated exactly. They correspond to a delta peak at $L = 0$. Another advantage over the approximation presented in Eq. (24) is that the approximation is performed for each number of defaults j separately and weighted by j/K accordingly. In this approximation, the omitted third term is of the order j/K^3 and thereby much smaller than the third term of the simple second order approximation (30), which would be of the order $1/K^2$. Thus, when approximating each term in the combinatorial sum separately, we obtain an improved approximation. Insertion into (31) leads to

$$\langle p(L) \rangle \approx \frac{1}{2\pi\Gamma(N/2)} \sum_{j=0}^{K} \binom{K}{j} \int_0^{\infty} dz \ z^{\left(\frac{N}{2}-1\right)} \exp\left(-z\right)$$
$$\left(\sqrt{\frac{2\pi K^2}{j\left(m_2(z) - m_1(z)^2\right)}} \exp\left(-\frac{(LK - j m_1(z))^2}{2j\left(m_2(z) - m_1(z)^2\right)} \right) \right.$$
$$\left. \times \left(\frac{1}{2} + \frac{1}{2} \mathrm{Erf}\left[\frac{\ln(F/v_0) - (\mu - \sigma^2/2)T}{2\sigma\sqrt{zT}} \right] \right)^{K-j} \right) , \tag{37}$$

which is the final result.

V. APPLICATION

We now apply the formally developed model to a specific example. To analyze the impact of correlations, we calculate the loss distribution for different homogenous portfolios with sizes $K = 10$, $K = 50$ and $K = 100$ with the parameters $v_0 = 100$, $\mu = 0.05$, $\sigma = 0.15$, $f = 75$ and $T = 1$. As stated in the previous section, we can control the

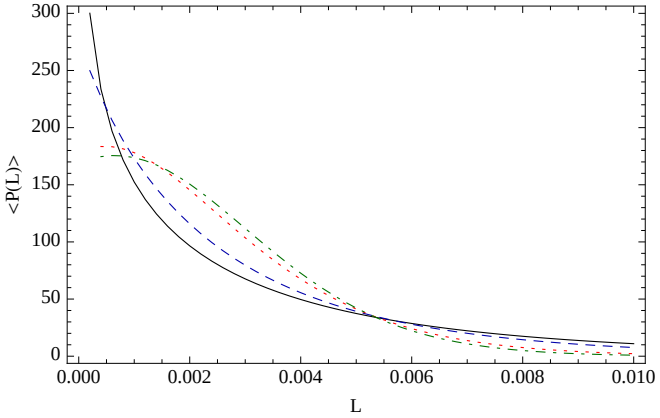


FIG. 3: The loss distribution for $K = 10$, $\sigma = 0.15$, $\mu = 0.05$, $T = 1$, $v_0 = 100$, $f = 75$ and differentamounts of randomness in the correlation matrix, $N = K$ (solid black), $N = 2K$ (dashed blue), $N = 10K$ (dotted red), $N = 30K$ (dot-dashed green).

amount of correlation in our model with the parameter N. To analyze the impact of random correlations, we choose $N = K$. For $N \to \infty$, the correlation matrix becomes the identity matrix. Thus, this represents the transition to a system without correlations. As we have to evaluate the loss distributions numerically, $N \to \infty$ is not practical. Hence we need to identify a value for which this convergence is valid in good approximation.

Fig. 3 illustrates exemplarily the loss distribution for $K = 10$ and different values of N. Our study indicates that a value of $N = 30K$ is a good choice for approximating the uncorrelated case and is still numerically feasible.

The results are presented in Fig. 4. For all portfolio sizes, $K = 10$, $K = 50$ and $K = 100$, we obtain heavier tails of the loss distribution of the correlated portfolio compared to the uncorrelated case. Even the second order approximation, represented by the dashed blue curve, exhibits these heavy tails. With the inserted logarithmic plots, we can identify a nearly power-law decay of the loss distribution for the correlated case.

The distributions become narrower for larger values of K. However, the tails of the correlated case remain heavier than those of the uncorrelated case. While for $K = 10$, both approximations give similar results, their difference becomes larger with K. As both approximations have to be performed numerically, we suggest to always employ the improved approximation. However, the tail behavior remains the same, even for the second order approximation, as indicated by the logarithmic scaled inserts in Fig. 4. This is a strong indication that the tails of the loss distribution are vastly underestimated, if correlations are not taken into account.

Due to the approximation, the normalization of the loss distribution is not exact. Especially the normalization of the second order approximation is poor for large values of K. The normalization might also be used as an indication for the quality of the normalization. The improved approximation exhibits a delta peak at $L = 0$, as the non-default terms can be calculated exactly. However, the interval $[0; 0.0002[$ was not evaluated due to numerical feasibility.

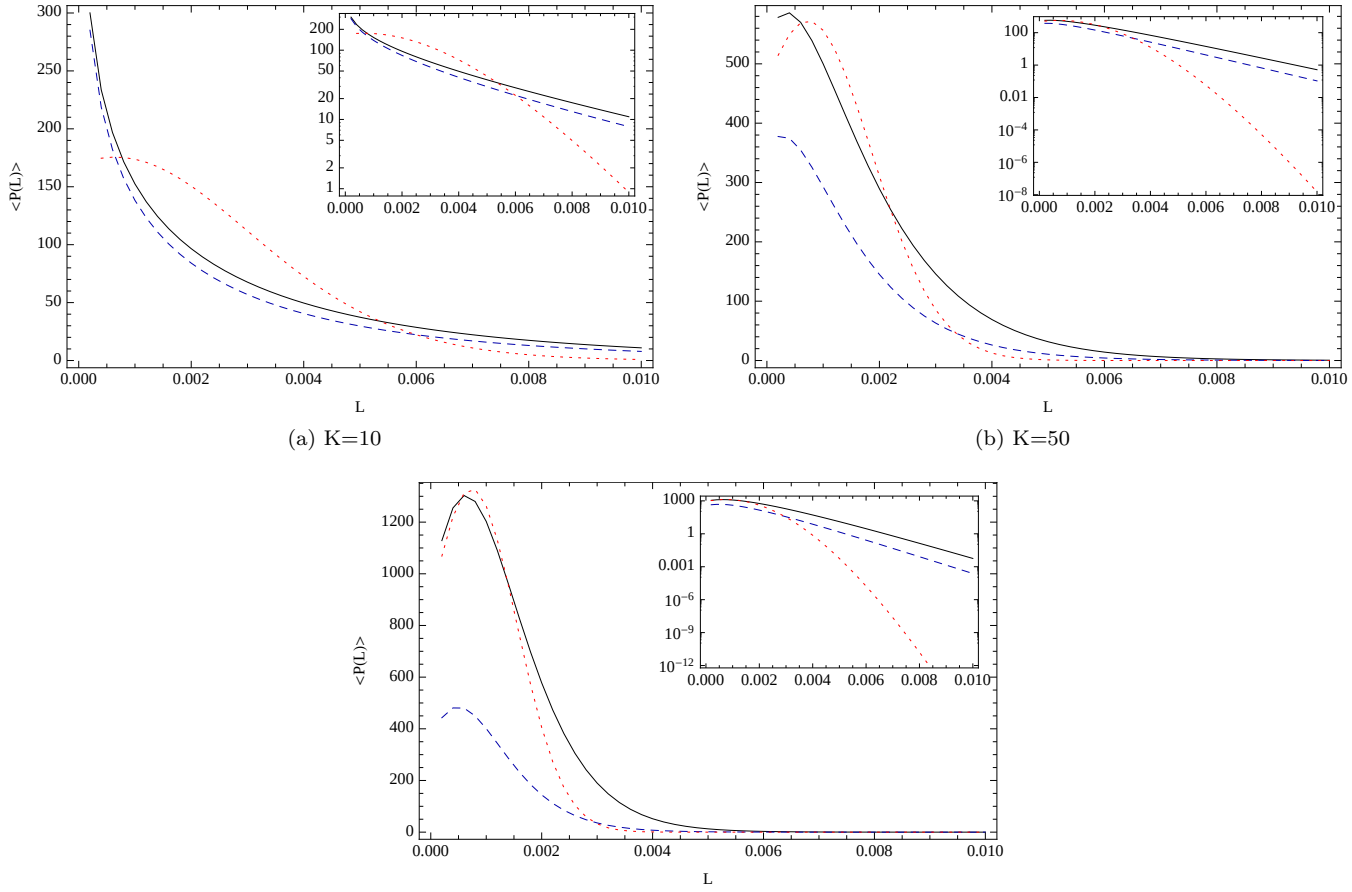
In our example, we do not vary the maturity time T, i.e., we choose $T = 1$. If the model is properly set up, one can increase T to estimate the evolution of the loss distribution. However, this evolution depends strongly on the drifts μ_k and standard deviations σ_k. Depending on their value, the exposure to default risk can either increase or decrease.

VI. SUMMARY

Our results clearly demonstrate that the risk in a credit portfolio is heavily underestimated if correlations are not taken into account in the quantification of risk. Even with random correlations and an average correlation level of zero, we obtain distributions whose tails differ in magnitude of several orders. In contrast, the probability of large losses in uncorrelated portfolios is significantly reduced within the model. This suggests that the effect of diversification, the reduction of risk in large portfolios, does not give good results in credit portfolios. Even the presence of random correlations around zero lowers the effect of diversification dramatically.

Our model can be used to estimate the lower bound of the risk embedded in a credit portfolio. Studies indicate that the amount of randomness in stock market return correlations enormously large [29]. Hence, despite its high level of abstraction, our model is directly applicable to speculative margin loans, i.e., credits that are based on an investor's stock portfolio. An overview of the model's input parameters is given in Tab. I.

Due to the level of abstraction, we omitted some features, that are present in empirical data, such as jumps or



(a) K=10

(b) K=50

(c) K=100

FIG. 4: The loss distribution of a homogenous portfolio with $\sigma = 0.15$, $\mu = 0.05$, $T = 1$, $v_0 = 100$, $f = 75$ and different values of K. The blue dashed line represents the simple approximation; the solid black line represents the improved approximation. Both have been calculated with maximum random correlations, $N = K$. The uncorrelated case is given by the red dotted line, calculated with the improved approximazation with $N = 30K$.

an overall positive correlation level. Those features are difficult to treat completely analytically. However, even in this simple setup, with random correlations and an average correlation level of zero, we obtain a heavy-tailed loss distribution.

The results are especially relevant for CDOs, bundles of credits that are traded on equity markets. CDOs are constructed in order to lower the overall risk. The components of a CDO can be exposed to large risks. It is often believed that the CDO has a significantly lower risk. We showed that this diversification only works well if the correlations in the credit portfolio are identical to zero. Even if the average correlation level is zero, but the individual correlations fluctuate, we obtain a heavy-tailed loss distribution. However, the average correlation level, for example, of stock returns is usually positive, corresponding to the market risk. A common approach is to remove this kind of risk by hedging. This corresponds to a shift in the correlation matrix by the size of the market correlation. The result is a correlation matrix with average correlation level of zero, as we assumed in our model. Thus, even if a portfolio manager eliminates the market risk, the probability of large losses is still significant. In credit portfolios, the risk of simultaneous defaults is not only caused by positive correlations, but also by their fluctuations.

ACKNOWLEDGEMENT

M.C.M. acknowledges financial support from Studienstiftung des deutschen Volkes.

Variable	Description	Unit
K	Number of assets	–
T	Time of maturity	a
σ_k	volatility of the k-th asset	$a^{-1/2}$
μ_k	drift of the k-th asset	a^{-1}
N	Parameter to control correlations, $N \to \infty$: uncorrelated limit	–
$v_{k,0}$	start price of the k-th asset	[currency]
F_k	face value of the k-th asset	[currency]

TABLE I: Input of the structural credit risk model.

REFERENCES

[1] C. Bluhm, L. Overbeck, and C. Wagner, *An Introduction to Credit Risk Modeling*. CRC Press, Taylor and Francis, 2002.
[2] T. R. Bielecki and M. Rutkowski, *Credit Risk: Modeling, Valuation and Hedging*. Springer, 2005.
[3] D. Duffie and K. J. Singleton, *Credit Risk: Pricing, Measurement, and Management*. Princeton University Press, 2003.
[4] D. Lando, *Credit Risk Modeling: Theory and Applications*. Princeton University Press, 2004.
[5] A. J. McNeil, R. Frey, and P. Embrechts, *Quantitative Risk Management: Concepts, Techniques, and Tools*. Princeton University Press, 2005.
[6] R. Schäfer, M. Sjölin, A. Sundin, M. Wolanski, and T. Guhr, "Credit risk–a structural model with jumps and correlations," *Physica A*, vol. 383, no. 2, pp. 533–569, 2007.
[7] J. Bessis, *Risk Management in Banking*. Wiley, 2001.
[8] A. Bangia, F. X. Diebold, A. Kronimus, C. Schagen, and T. Schuermann, "Ratings migration and the business cycle, with application to credit portfolio stress testing," *Journal of Banking & Finance*, vol. 26, no. 2-3, pp. 445–474, 2002.
[9] E. I. Altman, "The importance and subtlety of credit rating migration," *Journal of Banking & Finance*, vol. 22, no. 10-11, pp. 1231–1247, 1998.
[10] H. M. Markowitz, "Portfolio selection," *The Journal of Finance*, vol. 7, no. 1, pp. 77–91, 1952.
[11] K. Giesecke, *Credit Risk: Models and Management*, vol. 2, ch. Credit Risk Modeling and Valuation: An Introduction, pp. 487–525. Risk Books, 2004.
[12] F. Black and M. S. Scholes, "The pricing of options and corporate liabilities," *Journal of Political Economy*, vol. 81, no. 3, pp. 637–54, 1973.
[13] R. C. Merton, "On the Pricing of Corporate Debt: The Risk Structure of Interest Rates," *Journal of Finance*, vol. 29, no. 2, pp. 449–470, 1974.
[14] R. A. Jarrow and S. M. Turnbull, "Pricing derivatives on financial securities subject to default risk," *Journal of Finance*, vol. 50, pp. 53–86, 1995.
[15] R. A. Jarrow, D. Lando, and S. M. Turnbull, "A markov model for the term structure of credit risk spreads," *Review of Financial Studies*, vol. 10, no. 2, pp. 481–523, 1997.
[16] D. Duffie and K. Singleton, "Modeling the term structure of defaultable bonds," *Review of Financial Studies*, vol. 12, pp. 687–720, 1999.
[17] J. C. Hull and A. White, "Valuing credit default swaps I: No counterparty default risk," *Journal of Derivatives*, vol. 8, no. 1, pp. 29–40, 2000.
[18] P. J. Schönbucher, *Credit Derivatives Pricing Models*. New Jersey: John Wiley & Sons, 2003.
[19] F. Black and J. C. Cox, "Valuing Corporate Securities: Some Effects of Bond Indenture Provisions," *Journal of Finance*, vol. 31, no. 2, pp. 351–367, 1976.
[20] Y. A. Katz and N. V. Shokhirev, "Default risk modeling beyond the first-passage approximation: Extended black-cox model," *Physical Review E*, vol. 82, no. 1, p. 016116, 2010.
[21] A. Asvanunt and A. Staal, *The Corporate Default Probability model in Barclays Capital POINT platform (POINT CDP)*. Portfolio Modeling, Barclays Capital, April 2009.
[22] A. Asvanunt and A. Staal, *The POINT Conditional Recovery Rate (CRR) Model*. Portfolio Modeling, Barclays Capital, August 2009.
[23] B. Rosenow and R. Weissbach, "Modelling correlations in credit portfolio risk," *Journal of Risk Management in Financial Institutions*, vol. 3, no. 1, pp. 16–30, 2009.
[24] G. M. Gupton, C. C. Finger, and M. Bhatia, "Creditmetrics – technical document," tech. rep., Morgan Guaranty Trust Company, 1997.

[25] "Credit Portfolio View, Approach Document und User's Manual," tech. rep., McKinsey & Company, 1998.
[26] "Credit Risk+: A Credit Risk Management Framework," tech. rep., Credit Suisse First Boston (CSFB), 1997.
[27] F. W. J. Olver, *Asymptotics and Special Functions*, pp. 37–38. Academic Press, 1974.
[28] G. N. Watson, *A Treatise on the Theory of Bessel Functions*, p. 183. Cambridge University Press, 1944.
[29] L. Laloux, P. Cizeau, J.-P. Bouchaud, and M. Potters, "Noise dressing of financial correlation matrices," *Physical Review Letters*, vol. 83, pp. 1467–1470, Aug 1999.

APPENDIX

Appendix A: Gaussian integral in average price distribution

$$\langle p(\mathbf{v}(T))\rangle = \int p(W)p(\mathbf{v}(T), SWW^\dagger S)d[W] \tag{A1}$$

$$= \left(\frac{1}{\sqrt{2\pi}}\right)^{2K+KN} \sqrt{N}^{KN} \int \exp\left(-\frac{N}{2}\mathrm{tr}\left(WW^\dagger\right)\right) \times \int \exp\left(-i\omega^\dagger \mathbf{v}(T)\right) \exp\left(-\frac{T}{2}\omega^\dagger SWW^\dagger S\omega\right) d[\omega]d[W] \tag{A2}$$

$$= \left(\frac{1}{\sqrt{2\pi}}\right)^{2K+KN} \sqrt{N}^{KN} \int \exp\left(-i\omega^\dagger \mathbf{v}(T)\right) \times \int \exp\left(-\frac{N}{2}\mathrm{tr}\left(WW^\dagger\right)\right) \times \exp\Big(-\frac{T}{2}\overbrace{\omega^\dagger SWW^\dagger S\omega}^{=\mathrm{tr}(\omega^\dagger SWW^\dagger S\omega)}\Big) d[W]d[\omega]\ . \tag{A3}$$

In the last steps, we took advantage of the fact that the term $\omega^\dagger SWW^\dagger S\omega$ results in a scalar, which can evidently be written as its trace. As the trace is invariant in cyclic permutation, we can express this term as $\mathrm{tr}(WW^\dagger S\omega\omega^\dagger S)$. Hence, we write with N denoting the identity matrix,

$$\langle p(\mathbf{v}(T))\rangle = \left(\frac{1}{\sqrt{2\pi}}\right)^{2K+KN} \sqrt{N}^{KN} \int \exp\left(-i\omega^\dagger \mathbf{v}(T)\right) \times \int \exp\left(-\frac{1}{2}\mathrm{tr}(WW^\dagger(NI + TS\omega\omega^\dagger S))\right) d[W]d[\omega] \tag{A4}$$

$$= \left(\frac{1}{\sqrt{2\pi}}\right)^{2K+KN} \sqrt{N}^{KN} \int \exp\left(-i\omega^\dagger \mathbf{v}(T)\right) \times \int \exp\left(-\frac{1}{2}\sum_{N=1}^{N}(\mathbf{w}_n^\dagger(NI + TS\omega\omega^\dagger S)\mathbf{w}_n)\right) d[\mathbf{w}_n]d[\omega] \tag{A5}$$

$$= \left(\frac{1}{\sqrt{2\pi}}\right)^{2K+KN} \sqrt{N}^{KN} \int \exp\left(-i\omega^\dagger \mathbf{v}(T)\right) \times \left(\int \exp\left(-\frac{1}{2}\overbrace{\mathbf{w}^\dagger}^{\equiv x_i}\overbrace{(NI + TS\omega\omega^\dagger S)}^{\equiv A_{ij}}\overbrace{\mathbf{w}}^{\equiv x_j}\right)d[\mathbf{w}]\right)^N d[\omega]\ . \tag{A6}$$

The last step can be accomplished, as the components of W are independent identically distributed. Thus, we can simplify the integration over the matrix W to the integration over the vector $\mathbf{w} \in \mathbb{R}^K$ to the power of N. The integral over $d[\mathbf{w}]$ is simply a Gaussian integral, as indicated by x_i, x_j and A_{ij}. As $\mathbf{w}_t$ consists of K components, this gives an additional factor $\sqrt{2\pi}^{KN}$ and thus leads to

$$\langle p(\mathbf{v}(T))\rangle = \left(\frac{\sqrt{N}^N}{2\pi}\right)^K \int \exp\left(-i\omega^\dagger \mathbf{v}(T)\right) \frac{1}{\sqrt{\det(NI + TS\omega\omega^\dagger S)}^N} d[\omega]\ . \tag{A7}$$

Appendix B: Derivation of component-wise average price distribution

$$\langle p(\mathbf{v}(T))\rangle = \left(\frac{1}{2\pi}\right)^K \frac{1}{\Gamma(N/2)} \int \exp\left(-i\omega^\dagger \mathbf{v}(T)\right) \times \int_0^\infty z^{\left(\frac{N}{2}-1\right)} \exp\left(-(1+\frac{T}{N}\omega^\dagger SS\omega)z\right) dz\, d[\omega] \tag{B1}$$

$$= \left(\frac{1}{2\pi}\right)^K \frac{1}{\Gamma(N/2)} \int_0^\infty z^{\left(\frac{N}{2}-1\right)} \exp\left(-z\right) \times \int \exp\left(-i\omega^\dagger \mathbf{v}(T)\right) \exp\left(-\frac{T}{N}\sum_{k=1}^K \sigma_k^2\omega_k^2 z\right) d[\omega] dz \tag{B2}$$

$$= \left(\frac{1}{2\pi}\right)^K \frac{1}{\Gamma(N/2)} \int_0^\infty z^{\left(\frac{N}{2}-1\right)} \exp\left(-z\right) \times \prod_{k=1}^K \left[\int \exp\left(-i\omega_k v_k(T)\right) \exp\left(-\frac{T}{N}\sigma_k^2\omega_k^2 z\right) d\omega_k\right] dz \tag{B3}$$

$$= \left(\frac{1}{2\pi}\right)^K \frac{1}{\Gamma(N/2)} \int_0^\infty z^{\left(\frac{N}{2}-1\right)} \exp\left(-z\right) \prod_{k=1}^K \left[\frac{\sqrt{\pi n}}{\sqrt{zT}\sigma_k} \exp\left(-\frac{N v_k^2}{4Tz\sigma_k^2}\right)\right] dz \tag{B4}$$

$$= \left(\frac{1}{2\pi}\right)^K \frac{1}{\Gamma(N/2)} \left(\prod_{k=1}^K \frac{1}{\sigma_k}\right) \times \int_0^\infty z^{\left(\frac{N}{2}-1\right)} \exp\left(-z\right) \left(\frac{\pi n}{zT}\right)^{\frac{K}{2}} \exp\left(-\frac{N}{4Tz}\sum_{k=1}^K \frac{v_k^2}{\sigma_k^2}\right) \tag{B5}$$

Appendix C: Normalization of the average price distribution

We can demonstrate this by calculating the variance of the i-th price v_i,

$$\mathrm{var}(v_i(T)) = \int d[v]\ v_i(T)^2 \langle p(\mathbf{v}(T))\rangle\ . \tag{C1}$$

We can solve this integral by using hyperspherical corrdinates.

$$\rho \equiv \sqrt{\sum_{k=1}^K \frac{v_k^2}{\sigma_k^2}} \quad , \quad \frac{v_i}{\sigma_i} \equiv \rho \cos(\vartheta)\ . \tag{C2}$$

Now we can write the integral in Eq. (C1) as

$$\mathrm{var}(v_i(T)) = \left(\sqrt{\frac{N}{2\pi T}}\right)^K \frac{1}{\Gamma(N/2)} \left(\prod_{k=1}^K \frac{1}{\sigma_k}\right) 2^{1-\frac{N}{2}} \int_0^\infty d\rho \rho^{K-1} \times \int_0^\pi d\vartheta \sin(\vartheta)^{K-2} \sigma_i^2 \rho^2 \cos(\vartheta)^2 \left(\sqrt{\frac{N}{T}}\rho\right)^{\frac{N-K}{2}} \mathcal{K}_{\frac{K-N}{2}}\left(\sqrt{\frac{N}{T}}\rho\right) \int d\Omega_{k-2}\ , \tag{C3}$$

with

$$\int d\Omega_{K-2} = \frac{2\pi^{(K-1)/2}}{\Gamma((K-1)/2)} \prod_{k=1}^K \sigma_k\ , \tag{C4}$$

we obtain

$$\begin{aligned}\mathrm{var}(v_i(T)) =& 2\sigma_i^2 \frac{T}{N} \frac{\Gamma(N/2+1)}{\Gamma(N/2)} && \text{(C5)}\\ =& \sigma_i^2 T \ . && \text{(C6)}\end{aligned}$$

Thus, the variance of v_i only depends on the standard deviation σ_i and the time T.

Appendix D: second order approximation of average loss distribution

We rearranged the constants so that each term in $r(\omega, z)$ is normalized to unity. $r(\omega, z)$ can now be written as

$$\begin{aligned} r(\omega,z) = \exp\Bigg(\sum_{k=1}^{K} \ln\Bigg[\Bigg(\int_0^{F_k} dv_k \overbrace{\exp\left(i\omega f_k \left(1-\frac{v_k}{F_k}\right)\right)}^{q(\omega,F_k)} + \int_{F_k}^{\infty} dv_k \Bigg) \\ \times \frac{\sqrt{N}}{2\sigma_k v_k \sqrt{\pi z T}} \exp\left(-\frac{N(\ln(v_k/v_{k,0}) - (\mu-\sigma^2/2)T)^2}{4zT\sigma_k^2} \right) \Bigg] \Bigg) \ . \end{aligned} \tag{D1}$$

We write $q(\omega, F_k)$ as

$$q(\omega, F_k) = \sum_{m=0}^{\infty} \frac{(i\omega f_k)^m}{m!} \left(1 - \frac{v_k}{F_k}\right)^m \ . \tag{D2}$$

Due to the normalization of $\langle p(\mathbf{v}) \rangle$, after insertion into Eq. (D1) the non-default term and the integral over first term of Eq. (D2) become one. Thus, we can start the sum at $m = 1$ and obtain

$$r(\omega, z) = \exp\left(\sum_{k=1}^{K} \ln\left(1 + \sum_{m=1}^{\infty} \frac{(i\omega f_k)^m}{m!} m_{m,k}(z) \right) \right) \tag{D3}$$

with

$$m_{m,k}(z) = \frac{\sqrt{N}}{2\sigma_k\sqrt{\pi z T}} \int_0^{F_k} \frac{1}{v_k} \left(1 - \frac{v_k}{F_k}\right)^m \exp\left(-\frac{N(\ln(v_k/v_{k,0}) - (\mu-\sigma^2/2)T)^2}{4zT\sigma_k^2} \right) dv_k \ . \tag{D4}$$

The integrals in Eq. (D4) can be expressed with the generalized hypergeometrical function ${}_p\mathcal{F}_q$. However, this integral representation might be more intuitive. Moreover, for explicit $m = 1, 2$ the integrals can be calculated in a closed form, although this results in bulky expressions.

Expanding the logarithm and collecting all terms up to the second order in F yields

$$\begin{aligned} \langle p(L) \rangle \approx & \frac{1}{2\pi\Gamma(N/2)} \int_0^{\infty} dz \ z^{\left(\frac{N}{2}-1\right)} \exp(-z) \int_{-\infty}^{+\infty} d\omega \exp(-i\omega L) \\ & \times \exp\left(\sum_{k=1}^{K} \left(i\omega m_{1,k}(z) f_k - \frac{\omega^2 f_k^2}{2} (m_{2,k}(z) - m_{1,k}(z)^2) \right) \right) && \text{(D5)} \\ = & \frac{1}{2\pi\Gamma(N/2)} \int_0^{\infty} dz \ z^{\left(\frac{N}{2}-1\right)} \exp(-z) \int_{-\infty}^{+\infty} d\omega \\ & \times \exp\left(i\omega \left(\left[\sum_{k=1}^{K} f_k m_{1,k}(z) \right] - l \right) - \frac{\omega^2}{2} \left[\sum_{k=1}^{K} f_k^2 \left(m_{2,k}(z) - m_{1,k}(z)^2 \right) \right] \right) \ . && \text{(D6)} \end{aligned}$$

Now we can solve the $d\omega$ integral leading to equation 24.