



UvA-DARE (Digital Academic Repository)

A dimension reduction approach for loss valuation in credit risk modeling

He, J.; Khedher, A.; Spreij, P.

DOI

<https://arxiv.org/abs/2401.00085>
[10.1142/S2424786323500640](https://doi.org/10.1142/S2424786323500640)

Publication date

2024

Document Version

Submitted manuscript

Published in

International Journal of Financial Engineering

License

CC BY

[Link to publication](#)

Citation for published version (APA):

He, J., Khedher, A., & Spreij, P. (2024). A dimension reduction approach for loss valuation in credit risk modeling. *International Journal of Financial Engineering*, 11(1), Article 2350058. <https://doi.org/https://arxiv.org/abs/2401.00085>, <https://doi.org/10.1142/S2424786323500640>

General rights

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

UvA-DARE is a service provided by the library of the University of Amsterdam (<https://dare.uva.nl>)

A dimension reduction approach for loss valuation in credit risk modelling

Jian He *¹, Asma Khedher ¹, and Peter Spreij ^{1, 2}

¹Korteweg-de Vries Institute, University of Amsterdam, Amsterdam, The Netherlands

²IMAPP, Radboud University, Nijmegen, The Netherlands

January 2, 2024

Abstract

This paper addresses the “curse of dimensionality” in the loss valuation of credit risk models. A dimension reduction methodology based on the Bayesian filter and smoother is proposed. This methodology is designed to achieve a fast and accurate loss valuation algorithm in credit risk modelling, but it can also be extended to valuation models of other risk types. The proposed methodology is generic, robust and can easily be implemented. Moreover, the accuracy of the proposed methodology in the estimation of expected loss and value-at-risk is illustrated by numerical experiments. The results suggest that, compared to the currently most used PCA approach, the proposed methodology provides more accurate estimation of expected loss and value-at-risk of a loss distribution.

keywords: Bayesian filter, credit risk, loss valuation

2020 Mathematics Subject Classification: 62P05, 91G40

*Corresponding author: j.he2@uva.nl

Contents

1	Introduction	3
1.1	Problem description and background	3
1.2	Contribution	4
1.3	Organization of the paper	5
2	Set up and background	5
2.1	Overview of loss calculations in credit risk	5
2.2	Valuation model for the credit losses	9
2.3	Valuation grid for losses valuation	11
2.4	valuation grid pre-processing and loss simulation	12
3	A dimension reduction approach: Bayesian filter projection	13
3.1	State space model	14
3.2	Bayesian filter and smoother	14
3.2.1	Bayesian filter recursion	15
3.2.2	Bayesian smoother recursion	16
3.3	Bayesian filter/smoothing projection	17
3.4	Application examples	19
3.4.1	Credit Risk Economic Capital	20
3.4.2	Incremental Risk Charge	21
3.4.3	International Financial Reporting Standard 9	21
3.4.4	Market risk	22
3.4.5	Interest rate risk in the banking book	22
4	Numerical studies	23
4.1	Transition model	23
4.2	LGD model	25
4.3	Experiments	26
4.3.1	Transition probabilities	27
4.3.2	Expected PD simulation	27
4.3.3	Experiments for the projection approach for expected PD	28
4.3.4	Expected LGD: Black-Scholes formula v.s. Monte Carlo	30
4.3.5	Experiments for the projection approach for risk measures of the loss distribution	31
5	Conclusions	33
A	Analytic formula for expected loss-given-default	37

1 Introduction

1.1 Problem description and background

In this paper we apply *the grids methodology* to construct an approximation of *the valuation function* which is used to evaluate potential credit losses of loans. To tackle the “curse of dimensionality” problem that the grids methodology heavily suffers from, we propose a dimension reduction technique based on the Bayesian filter and smoother.

Computing risk metrics in credit risk for large portfolios is computationally intensive. This is especially true when the analytical loss distribution is not available. In general, Monte Carlo simulation is required to estimate the loss distribution. In such a case, one needs to simulate the multi-year default probabilities (PD) and realize loss-given default (LGD) of each single loan or each pool of loans, and then evaluate the potential losses. The simulation of the multi-year PD and realized LGD for large and complex portfolios requires extremely large computational resources. Therefore, in practice a valuation grid is usually used to approximate the portfolio valuation functions to avoid multi-year simulations. Valuation grids are a commonly applied computational tool in science and engineering (for example, Thompson et al. (1998)) and have also been applied in finance. For instance, Broadie and Glasserman (2004) uses grids to speed up the valuation of high dimensional American options. Recently, the grids method are also used together with the neural networks or Gaussian process regression to obtain better precision and faster calculations in valuations, see Horvath et al. (2021), Pagnottoni (2019), Evans and Nair (2018). The idea of using grids to measure portfolio risk is used in the stratified sampling methodology by Jamshidian and Zhu (1996) to approximate directly the probability density function of a portfolio value.

The grids methodology that applies to approximate the valuation function needs only a small number of exact valuations, which are used to construct a grid. The constructed grid is a fast approximation of the valuation function that can then be used during a simulation. The key to applying this method is to be able to express the valuation function as a function of a small number of risk factors, usually not larger than three, since the grids methodology heavily suffers from the “curse of dimensionality”, Bellman (1957). The “curse of dimensionality” describes the extraordinarily rapid growth of computation time as the number of the dimension increases. When the dimensionality of the risk factors is large, a large number of valuations is required to construct the interpolation pricing function and hence the grid methodology becomes impractical. Nevertheless, in credit risk modelling, a higher factor model is preferred to better describe the transition probabilities. Therefore, a dimension reduction approach is needed to map the higher dimensional factors to lower dimensional factors that can be used in the valuation grids. Currently, the most used dimension reduction approach is Principal Component Analysis (PCA), Jolliffe (1990). It transforms the original risk factors into a new risk factor space composed of orthogonal principal directions. For example, five dimensional risk factors result in five uncorrelated principal components. Movements of the five dimensional factors can thus be described by the principal components. The principal components are ordered from the largest to the least in terms of explaining the total variance of the five dimensional risk factors. In practice, one needs to decide on the number of principal components to use. Assume one decides to use the first two principal components. Then the original five dimensional risk factors are “reduced” to two dimensional risk factors, which are a linear combination of the first two principal components. However, using the PCA approach in credit risk modelling has the following drawbacks.

- *Low interpretability*: principal components are linear combinations of the risk factors, but they are not easy to interpret. For example, usually the risk factors need to be standardized before PCA decomposition. Consequently, it is difficult to tell which are the most important risk factors after computing principal components. Therefore one loses the economic interpretation when analyzing the impact of the principal components.
- *Sensitivity to outliers*: PCA is a method based on the correlation or covariance matrix, which can be very sensitive to outliers. Consequently, outliers can completely mess up a classical PCA analysis and yield a PCA model with an arbitrarily bad fit to the genuine part of the data. Indeed, PCA cannot account for outliers in the risk factors that are outside the range of the selected components and such outliers usually need to be removed before performing PCA. However, in credit risk modeling, these outliers are actually desired since they might be located in the tail of the loss distribution and hence contribute to the Value-at-Risk (VaR).
- *Assumption on the correlations*: PCA assumes correlations between the risk factors. If the risk factors are not correlated, PCA will not provide any additional insight.

As an alternative to the PCA approach, in this paper we propose a dimension reduction methodology.

1.2 Contribution

Our first and main contribution is a dimension reduction methodology based on the Bayesian filter and smoother. The Bayesian filter and smoother estimate the distribution (or value) of a latent process, given the observed data. We interpret the Bayesian filter and smoother as a projection from the observed data to unobserved latent factors. More precisely, this feature is applied to project the *higher* dimensional risk factors onto *lower* dimensional factors.

Compared to the commonly used Principal Components Analysis (PCA) approach, the proposed dimension reduction methodology has the follow advantages.

- *Economic interpretability*: The proposed approach projects the higher dimensional factors to lower dimensional factors in such a way that the lower dimensional factors capture the most importance features in the model maintaing an economic interpretation. For instance, if one applies the proposed approach to a transition model in which a certain rating migration constitutes the majority of observations (i.e. the majority of clients), this rating migration is an importance risk driver and the lower factors will tend to capture the feature of this rating migration.
- *Optimality*: The proposed approach is optimal in the sense that the Bayesian filter is optimal on the criterion of Bayes risk of minimum mean square error (MMSE), see Chen (2003). Intuitively, the proposed approach is optimal in that it seeks the posterior distribution of the lower dimensional factors which integrates and uses all of available information.
- *It is generic and robust*: This approach is generic in that it applies to a wide range of risk or pricing models, and insensitive to outliers.

Our second contribution is an application of our proposed dimension reduction approach to credit loss estimation. We review the important concepts and models used in the credit

risk modelling, like transitions, loss given default (LGD), exposure at default (EAD) and loss valuation models. We also derive a closed form formula of Black-Scholes type for the expected LGD of collateral LGD models. We conduct numerical experiments to illustrate the performances of the proposed methodology on the credit loss estimation and we assess the accuracy of the estimation on PD curves, expected losses, and the VaR at different levels. We compare the performance of the valuation grid based on the proposed methodology and the PCA approach. The results indicate that the proposed approach convincingly outperforms the PCA approach.

1.3 Organization of the paper

Section 2 provides the background knowledge for loss calculations in credit risk modelling. We present the valuation model and the valuation grid approach for the loss calculation. This section also contains brief reviews of transition, LGD and EAD models and specifies the models used in this paper. The proposed dimension reduction methodology is presented in Section 3, in which the Bayesian filter and smoother are also presented. In Section 4 we provide numerical experiments to show the performance of the proposed approach. Finally, Section 5 is devoted to the conclusions. In the Appendix we provide a Black-Scholes formula for the expected LGD.

Here follows some notations and other conventions used in this paper. All random processes are defined on a fixed probability space $(\Omega, \mathbb{F}, \mathbb{P})$. Time is assumed to be discrete. We will use a filtration $\{\mathcal{F}_k, k = 0, \dots, n\}$, where \mathcal{F}_k summarizes the information up to time k and n is the final time. We denote by \cdot^\top the transpose of a vector or a matrix.

2 Set up and background

In this section we outline the set up, we pose the problem, give a brief survey of models used for loss calculations in credit risk modeling, Bayesian filters (including Kalman, and particle filter) and address the ‘‘curse of dimensionality’’ problem for loss calculations using Bayesian filters.

2.1 Overview of loss calculations in credit risk

One of the main tasks in credit risk is to forecast and quantify future losses of a portfolio at a certain time horizon. The time horizon is a future time period over which the credit risk is assessed. For instance, in the regulations prescribed by the Basel Committee on Banking Supervision Basle Committee (1999) there is a one-year time horizon across all asset classes. The forecast risk can be quantified, for example, by expected losses or value at risk, depending on the purpose. In the International Financial Reporting Standards (IFRS) extensive reports of the expected losses for different obligors and sectors are required and for Economic Capital (EC) the extreme quantile of the loss distribution (usually the 99.95% quantile) needs to be calculated.

Three risk parameters are essential in the process of calculating the potential losses: the probability of default (PD), the loss given default (LGD) and the exposure at default (EAD). The PD is the probability that a borrower will fail to pay back a debt. The LGD is the fraction of a loan that is lost when a borrower defaults and the EAD is the predicted amount of the loss, the borrower may be exposed to, when a debtor defaults on a loan. They are

generally estimated using the available historical information and are assigned to operations and customers according to their particular characteristics. In this context, the credit rating tools (ratings and scorings) assess the risk for each transaction or customer according to their credit quality by assigning them a score. This score is then used in assigning risk metrics, together with additional information such as transaction seasoning, loan to value ratio, customer segment, etc. The increase in the number of default events in the current economic situation contributes to reinforce the soundness of the risk parameters by adjusting their estimates and by refining methodologies. The incorporation of data from the years of economic slowdown is particularly important for refining the analysis of the cyclical behavior of credit risk. The effect on the PD estimates and on the credit conversion factor is immediate.¹ An analysis of the impact on the LGD, however, requires waiting for the outcome of the recovery processes associated with the default events. Below we present a brief review to the models of the three risk parameters, i.e. PD, LGD and EAD.

Transition models

A model for the PD, or for transition matrices that govern rating migrations, depends on the rating system. A credit rating system uses a limited number of rating grades to rank borrowers according to their default probabilities. Ratings are assigned by rating agencies such as Fitch, Moody's and Standard & Poor's, but also by financial institutions themselves. Rating assignments can be based on a qualitative process or on default probabilities estimated with a scoring model. To translate default probability estimates into ratings, one defines a set of rating grades depending on the default probabilities. For example, borrowers are assigned grade AAA if their probability of default is lower than 0.02%, to grade AA if their probability of default is between 0.02% and 0.05% and so on. Once the ratings are defined, the concern is to determine the probability with which the credit risk rating of a borrower decreases or increases by a given degree and a transition (PD) model is required. These probabilities that a credit risk rating of a borrower decreases or increases, from one period to the next one, are usually collected in a transition matrix. Namely, the transition matrix measures the probabilities of the migrations between different credit ratings over specific time intervals. In the transition matrix, the element located at the i -th row and j -th column represents the probability of a borrower migrates from the i -th rating to the j -th rating. These are called transition probabilities and are often related to macroeconomic variables such as interest rates, inflation, gross domestic product (GDP), unemployment, etc. Alternatively, transition probabilities can also be modelled by using certain abstract latent processes.

The transition (PD) models in general can be divided into two main classes: structural and reduced form models.

Structural models link the up- or down-grade (or default) probabilities of a firm to the value of its assets and liabilities. The origin of structural credit risk models is the Merton model, Merton (1974), and then further developed by, for instance, Hull and White (2001), Avellaneda and Zhu (2001), Duffie and Singleton (2004), Albanese et al. (2003), and Jeanblanc and Rutkowski (1999). In these structural models, the latent credit process is called a credit index $(X_t)_{t \geq 0}$, reflecting unobservable credit quality over time which is driven by firm-specific variables such as the asset values. A default occurs if the credit index crosses a default barrier.

¹The credit conversion factor calculates the amount of a free credit line and other off-balance-sheet transactions (with the exception of derivatives) to an EAD amount and is an integral part in the European banking regulation since the Basel II accords, see Basle Committee (2006).

The dynamics of the credit index are specified by the model and the default barriers can be calibrated to historical migration or market data. Structural models can be generalized to describe not only defaults, but also rating migrations. This requires a mapping from the latent credit index to the rating states at each time, which corresponds to specifying an interval for the credit index that corresponds to each rating class. The intervals are defined by rating barriers, and rating changes occur when the credit index passes a barrier value.

Reduced form models, see for example, Jarrow and Turnbull (1995), Duffie and Singleton (2004), Frey and McNeil (2002), Wendin and McNeil (2006), Koopman et al. (2008), Jeanblanc and Li (2020), and Jeanblanc and Lecom (2008), give a description for the latent credit quality of an entity by assuming that it is driven by exogenous factors. They are based on a linear factor model for the latent variable, or equivalently, for the hazard rate of a firm. The key assumption of reduced form models is that the latent process or hazard rate process is driven by systematic external factors. For instance, in a reduced form model the latent credit process $(X_t)_{t \geq 0}$ can be decomposed into two independent parts: a systematic risk factor and an idiosyncratic risk factor. The systematic risk factor includes factors that are relevant for the business environment of the firms, such as the state of the economy, dummy variables for the industry sectors and regions, or other variables that are related to credit quality. The idiosyncratic risk factor is the firm-specific term. As in the structural model, in reduced form model a default occurs if the latent variable $(X_t)_{t \geq 0}$ falls below a certain threshold at a certain time. Same concept can be generalized to credit ratings by defining a mapping from the credit variable $(X_t)_{t \geq 0}$ to the rating states, by using threshold parameters for all ratings. These threshold parameters give the limits for the latent variable to move from current rating to other ratings, see for instance Nickell et al. (2000), Frey and McNeil (2002), Wei (2003), Duffie and Singleton (2004). Given the systematic factors, the default (transition) probabilities depends on the cumulative distribution function of the idiosyncratic factors. This distribution function is usually referred to as the response function. The most common choices for the response function are the probit function (standard normal distribution function) and the logit function. Compared to the structural model, the reduced form models are not explicitly based on the company's balance sheet. It only requires information generally available in financial markets and hence historical estimation methods can be easily used. But the consequence, and drawback, is that the models do not explain the economic reasons for default.

In this paper we use a reduced form approach to model the transition probability. We will describe this in more detail in Section 4.1.

LGD models

Loss Given Default (LGD) is one of the key determinants of the credit risks of loans and other credit exposures, which measures the severity of the loss when a debtor defaults. It is the share of an asset that is lost when a borrower defaults. An client may not end up with a loss if it is in default since it could be cured. Hence it is always worth noting that the LGD estimate is sensitive to the definition of default, which can substantially change the implied Recovery rate. Moreover, the LGD is client-specific because such losses are generally understood to be influenced by key transaction characteristics such as the presence or the value of a collateral. For example, in mortgages the houses themselves are usually the collateral. That means that if the borrower defaults, the bank could sell the house to compensate for the loss from missing payments from the borrower. Obviously, the LGD could be stochastic over time. Despite

its stochastic nature, statistical modeling of the LGD has been challenging in the academic literature and in banking practice. Therefore many credit risk models used in practice assume that LGD is a deterministic proportion of the exposures subject to impairment and ignore the fact that LGD can fluctuate according to the economic cycle. For example, Altman (2011) and Altman et al. (2005) show that default rates and recovery rates are strongly negatively correlated and measure a correlation of 0.75 between yearly average default rates and loss rates in the United States. They provide strong correlation evidence between macroeconomic growth variables (such as gross domestic product) and recovery rates and test the impact of correlated defaults and LGD. There are various models proposed to describe the LGD. Frye (2000a) and Frye (2000b) propose a structural model with a systematic risk factor representing the state of the economy that drives both defaults and LGD. The dependence of the default indicator and LGD on the common risk factor gives rise to a strong correlation between the two, which is in line with empirical evidence. Hillebrand (2006) introduces dependent LGD modeling into a multifactor latent variable framework, providing a good fit to corporate bond data. Marginal distributions for indicator functions and LGD can be specified. Frontczak and Rostek (2015) show that for the retail sector it is important to include collateral and suggested improvements for modelling LGD for this sector. We adopt this latter approach to model the LGD in this paper. The details are presented in Section 4.2.

EAD models

Exposure at Default (EAD) is the predicted amount of loss that may be exposed to when a debtor defaults on a loan. In many cases, such as residential mortgages and personal loans, the EAD are deterministic and can be simply taken from the current on-balance amount. For credit cards though, the revolving nature of the credit line poses challenges with regards to predicting the exposure at default time. As credit card customers may borrow more money (off balance) in the months prior to default, simply taking the current balance for non-defaulted customers would not produce a conservative enough estimate for the amount drawn by the time of default. Examples to model the EAD with an off balance amount can be found in Gumbo (2013) and Tong et al. (2016). In this paper, we treat the EAD as deterministic and equal to the on-balance amount. Finally, given the risk parameters we next explain how the portfolio losses are determined.

Portfolio losses

Given the PD, LGD and the EAD models, the realized losses of the obligors can be determined by simulating the risk factors and the migrations. If the simulated rating of an obligor becomes default at a certain time t , the loss of this obligor can be simply computed as $\text{LGD}_t \times \text{EAD}_t$, the notation should be obvious. After discounting and aggregating all the losses at different times, one can obtain the aggregated loss of the obligor. Then the portfolio loss for each realization is computed by summing the aggregated losses from all obligors.

By simulating a lot of scenarios, one obtains the distribution of the portfolio losses and hence determines the desired risk metrics, such as expected loss or value-at-risk. Figure 2.1 is presented as a simple illustration to determine the loss distribution of the portfolio. A number of scenarios are simulated until maturity. In this case the maturity is 30 periods and the horizon is one period. In each scenario, the migrations of all obligors are simulated over time and the losses at each time point are computed for the obligors that move to default at

that time point. Aggregating all the discounted losses from all time points for all obligors, one obtains the portfolio losses for this specific scenario. In the end the loss distribution can be obtained by repeating the calculations for all simulated scenarios. This procedure is very time

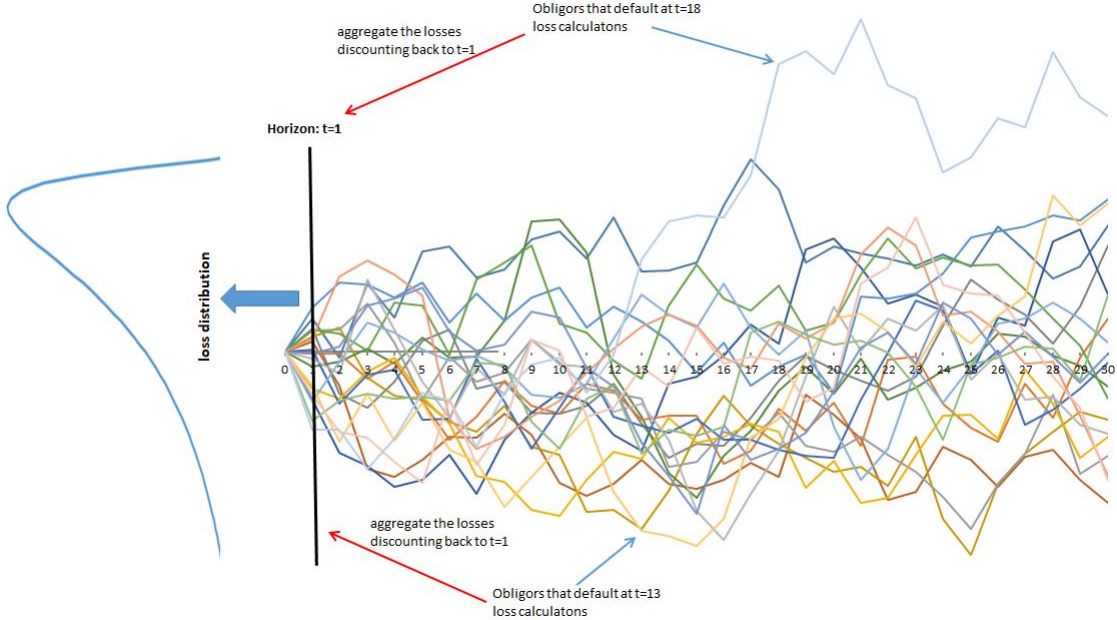


Figure 2.1: An illustration to determine the loss distribution

consuming when dealing with a huge portfolio, since it requires simulations of the migrations for all the obligors along each scenario. Hence, in practice, a valuation model together with a valuation grid method is needed to simplify the loss calculations. In the next subsections we introduce the valuation model we use in this paper and we introduce our valuation grid approach to compute the risk metrics.

2.2 Valuation model for the credit losses

In general, a portfolio's credit loss is defined as the difference between the portfolio's current value and its future value at the end of some time horizon see Basle Committee (1999). Therefore a valuation model can be used to compute the value of the loans and hence to capture the risks. For example, such a model is used to capture the risk of a one-year loan and a ten-year loan. If, over the one-year horizon, the one-year loan is downgraded but does not default, then it survives and there is no loss. If the ten-year loan is downgraded but does not default, then the *economic* value of this loan should be lower than if this loan was not downgraded. The loan values are based on a net present value (NPV). Specifically the NPV is computed as the expected discounted value of all future net cash flows:

$$\text{NPV} = \mathbb{E}\left(\sum_{k=1}^n \delta_k \text{CF}_k\right), \quad (2.1)$$

where n is the maturity of the loan, the $\delta_k \in \mathbb{R}^+$ are the risk-free discounting factors and the CF_k are the future cash flows at time k . It is very important to note, while it is correct

to use risk-free rates to discount risk-free cash flows, we have to discount risky cash flows due to potential defaults. This issue can be resolved in two ways: find risk-adjusted discount rates or transform the probabilities to risk adjusted (risk neutral) probabilities. The former way requires finding an appropriate discount rate for each (type of) obligor or loan, which would introduce a lot of parameters in the model. Therefore the latter approach to transform the physical default probabilities into risk-neutral default probabilities is usually used. More details about the risk neutral default probability can be found in Delianedis and Geske (2003). To reduce the complexity of the model in practice, the risk-free discounting factors are usually derived from deterministic curves, for instance those obtained from the Libor rates or Overnight Indexed Swaps (OIS). In this paper, we do not distinguish between the physical measure and the risk neutral measure for simplicity, and hence the expectation in equation (2.1) is interpreted as a risk neutral expectation.

Consider a loan with coupon payments s_k at time $k = 1, 2, \dots, n$ and a principal repayment PC at maturity n . Since the default depends on the rating of the loan, we denote the default time of a loan with rating r by τ_r , a non-negative real random variable. If the loan defaults in period k (between time $k - 1$ and k), denoted by the indicator $\mathbb{1}_{\tau_r \in (k-1, k]}$, the cash flow equals the coupon until time $k - 1$ plus the recovery $1 - \text{LGD}_k$. If the loan does not default, the cash flow equals all the coupons plus the principal. Hence, one obtains the value of the loan with rating r at initial time $k = 0$, denoted by $V_0^{(r)}$ as

$$\begin{aligned} V_0^{(r)} &= \mathbb{E} \left[\sum_{k=1}^n \mathbb{1}_{\tau_r \in (k-1, k]} \left(\sum_{j=1}^{k-1} \delta_j s_j + \delta_k (1 - \text{LGD}_k) \text{EAD}_k \right) + \mathbb{1}_{\tau_r > n} \left(\sum_{k=1}^n \delta_k s_k + \delta_n \text{PC} \right) \right] \\ &= \underbrace{\mathbb{E} \left[\sum_{k=1}^n \delta_k \mathbb{1}_{\tau_r \in (k-1, k]} (1 - \text{LGD}_k) \text{EAD}_k \right]}_{\text{lifetime PD}} + \underbrace{\mathbb{E} \left[\sum_{k=1}^n \delta_k \mathbb{1}_{\tau_r > k} s_k \right]}_{\text{coupon payments}} + \underbrace{\mathbb{E} [\delta_n \text{PC} \mathbb{1}_{\tau_r > n}]}_{\text{principle}}. \end{aligned}$$

Hence, standing at horizon h and suppose the sigma-algebra \mathcal{F}_h contains the information up to time h ,² the resulting value of a performing loan with rating r , denoted by $V_h^{(r)}$, is

$$\begin{aligned} V_h^{(r)} &= \text{RV}_h + \sum_{k=h+1}^n \delta_k \mathbb{E} [\mathbb{1}_{\tau_r \in (k-1, k]} (1 - \text{LGD}_k) \text{EAD}_k \mid \mathcal{F}_h] \\ &\quad + \sum_{k=h+1}^n \delta_k s_k \mathbb{E} [\mathbb{1}_{\tau_r > k} \mid \mathcal{F}_h] + \delta_n \text{PC} \mathbb{E} [\mathbb{1}_{\tau_r > n} \mid \mathcal{F}_h], \end{aligned} \tag{2.2}$$

where RV_h is the summation of the cash flow until horizon, i.e.

$$\text{RV}_h = \sum_{i=1}^h \delta_i s_i.$$

Note that $V_h^{(r)}$ is \mathcal{F}_h measurable.

²In practice, the specific definition of \mathcal{F}_h depends on the models used for PD, LGD, and EAD. For instance, if factor models are used, the sigma-algebra \mathcal{F}_h is usually defined as the sigma-algebra that is generated by the factors upto time h .

If there is no default after horizon h , the value of the loan is just the summation of all the discounted cash flows,

$$V_{h,\text{ND}}^{(r)} = \text{RV}_h + \sum_{k=h+1}^n \delta_k s_k + \delta_n \text{PC}.$$

Therefore, one obtains the valuation model for the loss of a performing loan with rating r at horizon h , conditioned on information at h , as follows.

$$\begin{aligned} L_h^{(r)} &= V_{h,\text{ND}}^{(r)} - V_h^{(r)} \\ &= \sum_{k=h+1}^n \delta_k s_k \mathbb{E}[\mathbb{1}_{\tau_r \leq k} | \mathcal{F}_h] + \delta_n \text{PC} \mathbb{E}[\mathbb{1}_{\tau_r \leq n} | \mathcal{F}_h] \\ &\quad - \sum_{k=h+1}^n \delta_k \text{EAD}_k \mathbb{E}[\mathbb{1}_{\tau_r \in (k-1, k]} (1 - \text{LGD}_k) | \mathcal{F}_h]. \end{aligned} \tag{2.3}$$

Moreover, note that in the last equality we consider the EAD_t as the on-balance amount at time t , as in most cases, and hence it is deterministic. One observes that the loss at time k is from the future cash flows after the defaults, but compensated by the recovery $(1 - \text{LGD}_k)\text{EAD}_k$. It is worth noticing that the definition of loss in this valuation model is in the *economic* sense. That means once a loan defaults, the losses not only come from the outstanding exposures at default, but also from missing interest payments from the future coupons. Note that the loss in Equation (2.3) is path-dependent. The probability of default depends on the value of the risk factors in the transition model and the LGD_k may also depend on risk factors in the LGD model.

2.3 Valuation grid for losses valuation

There is usually no closed-form expression for the conditional distribution of the cumulative PD or the conditional joint distribution of PD and LGD. Accordingly it is not possible to derive analytical formulas for the conditional expectations $\mathbb{E}[\mathbb{1}_{\tau_r \leq k} | \mathcal{F}_h]$ and $\mathbb{E}[\mathbb{1}_{\tau_r \in (k-1, k]} \text{LGD}_k | \mathcal{F}_h]$ in Equation (2.3). As alternative, a valuation grid is usually used to approximate the aforementioned conditional expectations. The valuation grid methodology, see for instance, Chishti (1999) and Gibson and Pritsker (2000), requires only a small number of valuations, which are used to construct a fast approximation of the valuation function (in our case the conditional expectations in (2.3)) when doing the simulation. The basic idea behind the valuation grid approach is simple. Consider a conditional expectation in which the value depends on a small number of risk factors (say two or three). Firstly, the conditional expectation values are evaluated, by Monte Carlo for instance, on a small, pre-selected set of points which forms a grid in the risk factor space. Then, during the simulation to construct the loss distribution, the conditional expectation values under different scenarios are quickly calculated by using the previously computed values on the grid points through interpolation or more sophisticated tools such as Gaussian process regression, see Rasmussen and Williams (2006) or neural networks, see Bishop (1994) and Angelini et al. (2008).

We apply the valuation grid to every conditional expectation in (2.3), that is to $\mathbb{E}[\mathbb{1}_{\tau_r \leq k} | \mathcal{F}_h]$ and to $\mathbb{E}[\mathbb{1}_{\tau_r \in (k-1, k]} \text{LGD}_k | \mathcal{F}_h]$ for $k = h + 1, \dots, n$. These conditional expectations are also called *bullets* since they can be seen as the loss of a unit bullet loan³. In the rest of

³A unit bullet loan is a loan with unit notional where a payment of the entire principal of the loan, possibly the principal and interest, is due at the end of the loan term.

our paper, we adopt this terminology and refer the conditional expectations as bullets. The key to applying the valuation grid method for valuation is to be able to express the bullet as a function of a small number of risk factors. However, these risk factor inputs are often high dimensional, since high dimensional factors are usually needed to describe the transition probabilities. The high dimensionality will result in big challenges in the implementation of the valuation grid, since the number of grid points to maintain a ‘minimum density’ grows too fast. This can be understood intuitively with a simple example. Suppose five risk factors are used to determine the value of a loan. If one wants to create a valuation grid with ten points in each of the five dimensions, then 10^5 exact valuations would be required to create the grid. To tackle this problem, we propose a dimension reduction approach to project a higher factor model onto a lower factor model for the valuation. The advantage of the proposed projection approach is that the higher factor model is kept to model the transition probabilities, whereas a corresponding lower factor model is used for the valuation. Further details of the proposed projection approach will be presented in Section 3.

2.4 valuation grid pre-processing and loss simulation

Suppose factor models are assigned to model the transitions and LGDs. Denote the risk factors at time k by x_k . The filtration $\{\mathcal{F}_s, s = 0, 1, \dots, n\}$ is the information set in which \mathcal{F}_s is the sigma-algebra generated by $\{x_0, \dots, x_s\}$. For every time k , every initial rating and possibly every initial loan-to-value (depending on the LGD model), the values of the bullets $\mathbb{E}[\mathbb{1}_{\tau_r \leq k} | \mathcal{F}_h]$ and $\mathbb{E}[\mathbb{1}_{\tau_r \in (k-1, k]} \text{LGD}_k | \mathcal{F}_h]$ in Equation (2.3) are pre-computed on certain pre-selected points of risk factors. Then given time k , initial rating r and possibly initial loan-to-value, the set of bullet values and the risk factor points form a valuation grid. The pre-computation of the bullets are usually done by Monte Carlo simulation.

Suppose the default indicator $\mathbb{1}_{\tau_r \in (k-1, k]}$ and LGD_k are conditionally independent given \mathcal{F}_k . Then one derives

$$\begin{aligned} \mathbb{E}[\mathbb{1}_{\tau_r \in (k-1, k]} \text{LGD}_k | \mathcal{F}_0] &= \mathbb{E}[\mathbb{E}[\mathbb{1}_{\tau_r \in (k-1, k]} \text{LGD}_k | \mathcal{F}_k] | \mathcal{F}_0] \\ &= \mathbb{E}[\mathbb{E}[\mathbb{1}_{\tau_r \in (k-1, k]} | \mathcal{F}_k] \mathbb{E}[\text{LGD}_k | \mathcal{F}_k] | \mathcal{F}_0] \\ &= \mathbb{E}[\mathbb{P}(\tau_r \in (k-1, k] | \mathcal{F}_k) \mathbb{E}[\text{LGD}_k | \mathcal{F}_k] | \mathcal{F}_0] \\ &= \mathbb{E}[(\text{PD}_r(k | \mathcal{F}_k) - \text{PD}(k-1 | \mathcal{F}_k)) \mathbb{E}[\text{LGD}_k | \mathcal{F}_k] | \mathcal{F}_0], \end{aligned}$$

with $\text{PD}_r(k | \mathcal{F}_k)$ the cumulative PD of loan with initial rating r , at time k conditioned on the information \mathcal{F}_k . This result suggests that, in the case that $\mathbb{1}_{\tau_r \in (k-1, k]}$ and LGD_k are conditionally independent given \mathcal{F}_k , one only needs to simulate transition (default) probabilities and expected LGD when using Monte Carlo simulations to estimate $\mathbb{E}[\mathbb{1}_{\tau_r \in (k-1, k]} \text{LGD}_k | \mathcal{F}_0]$. The additional simulations of the indicator $\mathbb{1}_{\tau_r \in (k-1, k]}$ or the realized LGD_t are not necessary. Therefore, the valuation function (2.3) at horizon h can be reformulated as

$$\begin{aligned} L_h^{(r)} &= \sum_{k=h+1}^n \delta_k s_k \mathbb{E}[\text{PD}(k | \mathcal{F}_k) | \mathcal{F}_h] + \delta_n \text{PCIE}[\text{PD}(n | \mathcal{F}_n) | \mathcal{F}_h] \\ &\quad - \sum_{k=h+1}^n \delta_k \text{EAD}_k \mathbb{E}[(\text{PD}(k | \mathcal{F}_k) - \text{PD}(k-1 | \mathcal{F}_k)) (1 - \text{LGD}_k) | \mathcal{F}_h]. \end{aligned} \tag{2.4}$$

When constructing the loss distribution, firstly the paths of the risk factors need to be simulated until the horizon. Then based on the simulated risk factors, the rating transitions and

the LGDs can be simulated (computed) according to the transition and LGD models. If the loan defaults or expires before the horizon, the losses are already obtained by the simulation. If the loan is still performing (not defaulted or expired) at the horizon, the valuation grids are then used to value the bullets and consequently the loss after the horizon according to the valuation model. Aggregating the losses of all the loans, one obtains the distribution of the aggregated loss. In Figure 2.2 a simple illustration is presented that shows how the valuation grid is used to approximate the loss distribution. Scenarios and migrations are simulated only until the time horizon. Then given each simulation, the realised loss is computed for each loan. Aggregating the losses from all the loans, the loss distribution is obtained by collecting the portfolio losses from all the scenarios.

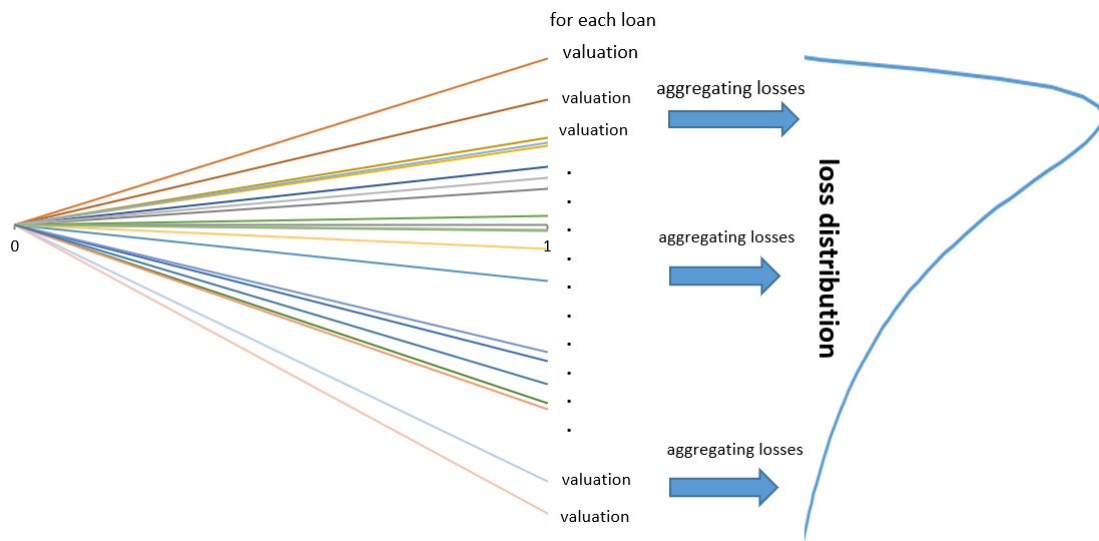


Figure 2.2: An illustration to determine the loss distribution using valuation grids

As stated, the losses happened before or on the horizon can be directly computed using the simulated transitions and LGDs and only the losses that happened after the horizon will be evaluated by the valuation grid. Therefore, in this paper we ignore the pre-horizon losses and only focus on the losses after the horizon.

3 A dimension reduction approach: Bayesian filter projection

This section on our projection method is the core of the paper and the approach is motivated as follows. The desired advantage of the proposed logit transition model, see (4.2) and (4.4), is that it allows a high dimensional multi-factor model which describes the transition data accurately. However, the challenge of implementing a high dimensional transition model in credit risk modeling is the computational complexity in the expected loss valuation using valuation grid. Ideally, the high dimensional transition model is used to simulate the risk factors and determine the transitions, while lower dimensional factors are used for the valuation grid. Note that the simulated higher dimensional factors that are used to determine the transitions and the lower dimensional factors used in the valuation grid need to be consistent. Therefore a (dimension reduction) projection approach is required to map the higher dimensional factors to lower dimensional factors. The most used dimension reduction method approach to tackle

this problem is Principal Component Analysis (PCA). But this approach is not optimal to capture the transition probabilities in a lower dimensional space. Hence the performance of this approach is not assured, see the experiments in Sections 4.3.1, 4.3.3 and 4.3.5. In this section, we propose a dimension reduction method, which link the higher dimensional factors with the lower dimensional factor through the Bayesian filter (smoother). Since the Bayesian filter (smoother) achieves minimum mean square error (MMSE) Chen (2003), the proposed dimension reduction method is optimal for the criterion of the Bayes risk, i.e. MMSE, to project the simulated higher dimensional factors to lower dimensional factors. The Bayesian filter is a general probabilistic approach for estimating the (distribution) of the latent states in a state space model. The state space model and the Bayesian filter will be described in Sections 3.1 and 3.2 respectively. The proposed dimension reduction algorithm is described in Section 3.3.

3.1 State space model

State-space models deal with dynamic time series problems that involve unobserved variables or parameters that describe the evolution in the state of the underlying system. The general *state space* model, defined on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$, is as follows

$$\begin{aligned} x_k &= f_k(x_{k-1}, u_k), \quad x_0 \\ y_k &= h_k(x_k, v_k), \quad k \in \mathbb{N}^+, \end{aligned} \tag{3.1}$$

where $f_k : \mathbb{R}^d \times \mathbb{R}^p \rightarrow \mathbb{R}^d$, $h_k : \mathbb{R}^d \times \mathbb{R}^q \rightarrow \mathbb{R}^m$ are given Borel measurable functions, $\{u_k\}_{k \in \mathbb{N}^+}$ are p -dimensional and $\{v_k\}_{k \in \mathbb{N}^+}$ are q -dimensional white noise processes both independent of the initial condition x_0 , and mutually independent as well. Parameters in the functions f_k and h_k , together with the covariance of u_k and v_k can be seen as the parameters of the state space model, to which we collectively refer to as ψ .

We are interested in estimating the (latent) state process $\{x_k\}_{k \in \mathbb{N}^+}$, but only have access to the process $\{y_k\}_{k \in \mathbb{N}^+}$ which represents the observations. Because of the existence of the white noise in the data, estimating the values, or even the distributions, of the latent states $\{x_k\}_{k \in \mathbb{N}^+}$ by the observations $\{y_k\}_{k \in \mathbb{N}^+}$ is not trivial. There are different methodologies in the literature to estimate the latent process, see e.g. Press (2003); Chui and Chen (2017); Arulampalam et al. (2002), and the Bayesian filter (smoother) is one of the mostly used approach in practice. Especially, the Bayesian filter (smoother) can be considered as a projection from the observation space to the state space. The Bayesian filter and smoother will be explained in Section 3.2.

The state space model is widely used in many different areas such as economics and finance, statistics, computer science, and electrical engineering and neuroscience. Particularly in finance, it can be used to identify the latent variables including business cycles, expectations of certain economic variables, permanent income streams, reservation wages, etc. The state space model is also often applied to factor models to describe, for instance, interest rate and credit risk driving factors.

3.2 Bayesian filter and smoother

The *Bayesian filter* and *Bayesian smoother*, see e.g. Särkkä (2013); Press (2003); Robert (2007) for an overview, are used to estimate the distribution of the latent states $\{x_k\}_{k \in \mathbb{N}^+}$

in (3.1) given the parameters, which we denote ψ . Often, but not always, we emphasize the dependence of densities on the parameters in the notation.

We let the initial density function of x_0 be p_0 . The methodology in Bayesian filtering consists of two parts: prediction and update. At every time point k , the prediction part computes (estimates) the *prior distribution (density)* of x_k given the past observations up to time $k - 1$,

$$p(x_k \mid y_{1:k-1}, \psi),$$

and the update part computes (estimates) the *posterior distribution (density)* of x_k given the past up to time k ,

$$p(x_k \mid y_{1:k}, \psi).$$

The purpose of the Bayesian smoother⁴ is to compute the marginal posterior distribution of the state x_k at the time step k after receiving the measurements up to a time step T with $T > k$,

$$p(x_k \mid y_{1:T}, \psi).$$

The difference between filters and smoothers is that the filter uses the information from the previous and current period to estimate the posterior distribution of current state x_k , while the smoothers uses also the observations after the current period and hence produces optimal estimations.

It follows that the state space model (3.1) satisfies the properties of a stochastic system, i.e. at every (present) time $k \geq 1$ the future states and future observations (x_j, y_j) , $j \geq k$, are conditionally independent from the past states and observations (x_j, y_{j-1}) , $j \leq k$, given the present state x_k , see van Schuppen (1989). It then follows that $\{x_k\}_{k \in \mathbb{N}}$ is a Markov process, and for every $k \geq 1$ one has that y_k and $y_{1:k-1}$ are conditionally independent given x_{k-1} , in terms of densities,

$$p(x_k \mid x_{k-1}, y_{1:k-1}) = p(x_k \mid x_{k-1}). \quad (3.2)$$

Similarly, due to Markov property, x_k and $y_{k+1:T}$ are independent given x_{k+1} , which gives

$$p(x_k \mid x_{k+1}, y_{1:T}) = p(x_k \mid x_{k+1}, y_{1:k}). \quad (3.3)$$

Moreover, one also has, for every $k \geq 1$, that y_k and $y_{1:k-1}$ are conditionally independent given x_{k-1} , in terms of densities,

$$p(y_k \mid y_{1:k-1}, x_k) = p(y_k \mid x_k), \quad \text{for } k \in \mathbb{N}^+. \quad (3.4)$$

These three equations, (3.2), (3.3) and (3.4) are used in the next sections in the derivation of Bayesian filter and smoother recursions.

3.2.1 Bayesian filter recursion

Using Bayes' rule and Equation (3.2), we deduce that the density function of the prior distribution is given by

$$p(x_k \mid y_{1:k-1}, \psi) = \int p(x_k \mid x_{k-1}, y_{1:k-1}, \psi) p(x_{k-1} \mid y_{1:k-1}, \psi) dx_{k-1}$$

⁴This definition actually applies to the fixed-interval type of smoothing. For fixed-leg smoothing, one can consult Anderson and Moore (2012).

$$= \int p(x_k | x_{k-1}, \psi) p(x_{k-1} | y_{1:k-1}, \psi) dx_{k-1}. \quad (3.5)$$

Note that when $k = 1$, the posterior distribution $p(x_{k-1} | y_{1:k-1}, \psi)$ is defined as the initial density p_0 .

The purpose of the Bayesian algorithm is to sequentially compute the posterior distribution $p_k(dx_k | \psi) = p(dx_k | y_{1:k}, \psi)$. Again using Bayes' rule, (3.4) and (3.5), we obtain the posterior

$$\begin{aligned} p(x_k | y_{1:k}, \psi) &= \frac{p(y_k | x_k, y_{1:k-1}, \psi) p(x_k | y_{1:k-1}, \psi)}{p(y_k | y_{1:k-1}, \psi)} \\ &= \frac{p(y_k | x_k, \psi) p(x_k | y_{1:k-1}, \psi)}{\int p(y_k | x_k, \psi) p(x_k | y_{1:k-1}, \psi) dx_k}. \end{aligned} \quad (3.6)$$

If we assume the marginal likelihood function $p(y_k | x_k, \psi)$ and the transition probability $p(x_k | x_{k-1}, \psi)$ are known, then given the posterior density $p(x_{k-1} | y_{1:k-1}, \psi)$ at time $k - 1$, we can use Equations (3.5) and (3.6) to compute the posterior density $p(x_k | y_{1:k}, \psi)$ at time k . In this way the posterior distributions can be computed recursively given the initial distribution p_0 .

3.2.2 Bayesian smoother recursion

According to Bayes's rule and Equations (3.2) and (3.3), the density $p(x_k | y_{1:T})$ can be expressed as

$$\begin{aligned} p(x_k | y_{1:T}, \psi) &= \int p(x_k | x_{k+1}, y_{1:T}, \psi) p(x_{k+1} | y_{1:T}, \psi) dx_{k+1} \\ &= \int p(x_k | x_{k+1}, y_{1:k}, \psi) p(x_{k+1} | y_{1:T}, \psi) dx_{k+1} \\ &= \int \frac{p(x_k, x_{k+1} | y_{1:k}, \psi)}{p(x_{k+1} | y_{1:k}, \psi)} p(x_{k+1} | y_{1:T}, \psi) dx_{k+1} \\ &= \int \frac{p(x_{k+1} | x_k, y_{1:k}, \psi) p(x_k | y_{1:k}, \psi)}{p(x_{k+1} | y_{1:k}, \psi)} p(x_{k+1} | y_{1:T}, \psi) dx_{k+1} \\ &= p(x_k | y_{1:k}, \psi) \int \frac{p(x_{k+1} | x_k, \psi)}{p(x_{k+1} | y_{1:k}, \psi)} p(x_{k+1} | y_{1:T}, \psi) dx_{k+1}. \end{aligned} \quad (3.7)$$

According to Equation (3.7), given that the prior and posterior distribution are known from the Bayesian filter recursions, the Bayesian smoother estimation $p(x_k | y_{1:T}, \psi)$ can be computed using a backward recursion.

While the mathematical derivations of the Bayesian filter and smoother are straightforward, in practice it is always a big challenge to compute the integrals in Equations (3.5), (3.6) or (3.7). Although there are some special cases (Gaussian linear cases) where theoretical formulas are available for the integrals, such as Kalman filter (see Welch and Bishop (1995) for instance), in most cases one has to find approximations or numerical algorithms to compute these integrals. For the extended Kalman filter, see, for example, Eincike and White (1999) and Wan and Nelson (2002) and for the unscented Kalman filter, see Wan and van der Merwe (2002), the prior and the posterior distributions are approximated by Gaussian distributions.

Durbin and Koopman (2012, Section 10.6–10.7) proposes a mode estimation approach to compute the maximum-a-posterior (MAP) estimate of the posterior distribution. A more general approach to estimate the prior and posterior distribution is the particle filter, see e.g. Arulampalam et al. (2002); Cappé et al. (2007); Doucet and Johansen (2011), in which prior and posterior distributions are estimated by a Monte Carlo method. For the implementation of the particle filter one can also refer to Chopin et al. (2013); Crisan and Míguez (2018); He et al. (2021).

3.3 Bayesian filter/smoothing projection

As described in Section 3.2, the Bayesian filter and smoother can be seen as a projection from the observations $y_{1:k}$ to the (distribution of) state variables, given the model parameters ψ . From this perspective, the Bayesian filter (smoother) can be used as an approach to project higher dimensional factors to lower dimensional factors. In this section we will illustrate the Bayesian filter (smoother) projection approach in the credit risk environment specified in Sections 2.2. However, note that this approach can also easily be applied or extended to other models for different risk types, such as market risk or interest rate risk, see also the examples in Section 3.4.

Suppose transition models, see Section 2.1, with higher and lower dimension are calibrated based on the same data, and denote the calibrated parameters of higher and lower dimensional model by ψ_H and ψ_L , respectively. The higher dimensional factor model is used to simulate the transition probabilities until the horizon according to the transition model, for instance, a simulation based on Equations (4.4) and (4.2) in Section 4. Denote the simulated higher dimensional factors by $x_{k,H}$ and the corresponding transition probabilities by $T(x_{k,H})$, for $k = 1, \dots, h + 1$. One would like to use a lower factor model for a loss calculation. It is then required to project the higher dimensional factors $x_{k,H}$ to the lower dimensional factors, denoted by $x_{k,L}$. Note that the simulated transition probabilities $T(x_{k,H})$, based on the higher dimensional factors $x_{k,H}$, can be used as the observations of the lower factor model. Therefore, using the Bayesian filter and smoother iterations, one can obtain a projection from $x_{k,H}$ to $x_{k,L}$, as described in Algorithm 3.1.

Algorithm 3.1 (projection). Given are the calibrated parameters of the higher and lower factor model ψ_H and ψ_L . Suppose $x_{1:h+1,H}$ and $T(x_{1:h+1,H})$ are the simulated higher dimensional factors and the corresponding transition probabilities from time $k = 1$ to horizon $h + 1$.

1. Specify the initial density $p(x_0 | \psi_L)$ for the lower dimensional factors.
2. Using the transition probabilities $T(x_{1:h+1,H})$ as the data,
 - a. for $k = 1, \dots, h + 1$, compute the prior and posterior distribution $x_{k,L}$ according to the Bayesian filter iterations, see Section 3.2.1, i.e.

$$p(x_k | T(x_{1:k-1,H}), \psi_L) = \int p(x_k | x_{k-1}, \psi_L) p(x_{k-1} | T(x_{1:k-1,H}), \psi_L) dx_{k-1}$$

$$p(x_k | T(x_{1:k,H}), \psi_L) = \frac{p(T(x_{k,H}) | x_k, \psi_L) p(x_k | T(x_{1:k-1,H}), \psi_L)}{\int p(T(x_{k,H}) | x_k, \psi_L) p(x_k | T(x_{1:k-1,H}), \psi_L) dx_k};$$

- b. given prior and posterior distributions computed in step 2.a, for $k = h, \dots, 1$, compute the Bayesian smoother using the backward recursion, see Section 3.2.2,

$$p(x_k | T(x_{1:h+1,H}), \psi_L) = p(x_k | T(x_{1:k,H}), \psi_L) \times \int \frac{p(x_{k+1} | x_k, \psi_L)}{p(x_{k+1} | T(x_{1:k,H}), \psi_L)} p(x_{k+1} | T(x_{1:h+1,H}), \psi_L) dx_{k+1};$$

- c. determine the lower dimensional factor $x_{k,L}$ projected from $x_{k,H}$ as equal to

$$\mathbb{E}[p(x_k | T(x_{1:h+1,H}), \psi_L)], \quad k = 1, \dots, h.$$

We are in particular interested in the value of the lower dimensional factor at horizon h , i.e. $x_{h,L}$, since it is finally used as the input of the valuation grid to evaluate the losses after this horizon. Note that in order to obtain the optimal estimate of $x_{h,L}$, the simulation is generated until $h + 1$ so that the Bayesian smoother, step 2.b in Algorithm 3.1, can be used for $x_{h,L}$. Figure 3.1 presents the diagram of the projection approach.

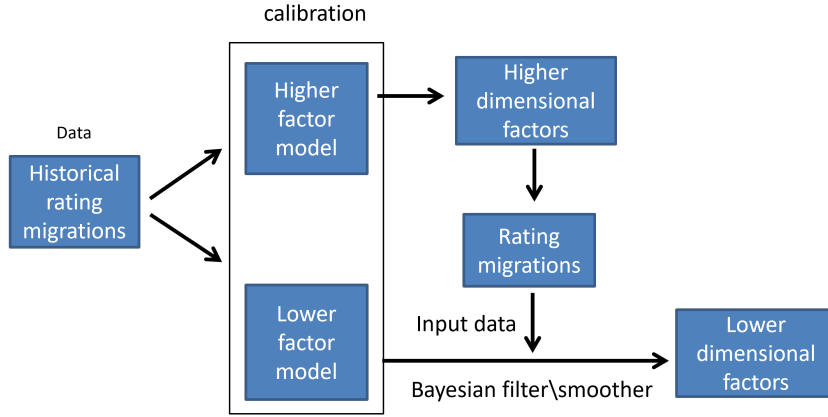


Figure 3.1: High level description of projection algorithm

Remark 3.1. An important input for the iteration of the projection is the initial distribution of $x_{k,L}$, i.e. $p(x_0 | \psi_L)$. In credit risk, there are two methodologies to specify the initial distribution (or starting point) of the simulation, the point-in-time (PiT) or through-the-cycle (TtC). In PiT, all the simulations start at the same point and this point is defined as the ‘state of today’, i.e. the value of the risk factor at the last period used in the calibration. In TtC, instead of defining a single starting point, an initial distribution is assigned to sample the starting point of each path. This initial distribution is usually defined as the equilibrium distribution of the risk factors process.⁵ Therefore, in both PiT and TtC cases, the initial distribution of $x_{k,L}$ can be easily defined.

Integrating the projection algorithm into the calibration, simulation and valuation algorithms, we have our engine for the risk calculation, illustrated by Figure 3.2, and described in the following Algorithm 3.2, formulated in high level terms.

⁵In practice, constraints are always imposed to make sure that the risk factor process is stationary and hence the equilibrium distribution exists.

Algorithm 3.2 (calibration, training, simulation, projection and valuation).

- **Calibration**

1. The transition matrices are modeled by a higher factor model, which is calibrated using historical data. One obtains ψ_H , the calibrated parameters of the higher dimensional model.
2. A lower factor model is also calibrated using the historical data and one obtains ψ_L , the calibrated parameters of the lower dimensional model.

- **Training of the interpolation grid**

Define the higher dimensional grid, perform then for each rating the following steps.

1. Starting at the points on the defined grid, use the higher factor model to simulate the scenarios of the risk factors and determine the expected losses.
2. Project the points (factors) of the higher dimensional grid to obtain lower dimensional points (factors) using the projection approach as in Algorithm 3.1.
3. Use the lower dimensional points and the simulated expected losses to build the valuation grid, using interpolation or other techniques like Gaussian process regression, Neural networks, etc.

- **Simulation**

1. Using the high factor model, generate the factors and hence the transition matrices until the horizon.
2. Based on the generated transition matrices, simulate the rating migrations of the obligors until the horizon.

- **Projection and valuation**

1. Project according to the Projection Algorithm 3.1 the higher dimensional factors simulated in Step 2 of the Simulation part above to lower dimensional factors.
2. Use the valuation grid to evaluate the expected losses, given the ratings at the horizon and the lower dimensional factors.
3. The desired risk metrics, such as expected loss and value-at-risk, can finally be obtained by aggregating the expected losses.

3.4 Application examples

The proposed Bayesian projection approach is a generic dimension reduction method which can be used for different types of credit risk valuations. It is originally designed for the credit risk valuations, but it can also be applied to other grid based valuation models where the dimension reduction approach is needed. We list some possible directions for further applications of the proposed Bayesian projection approach in the following subsections.

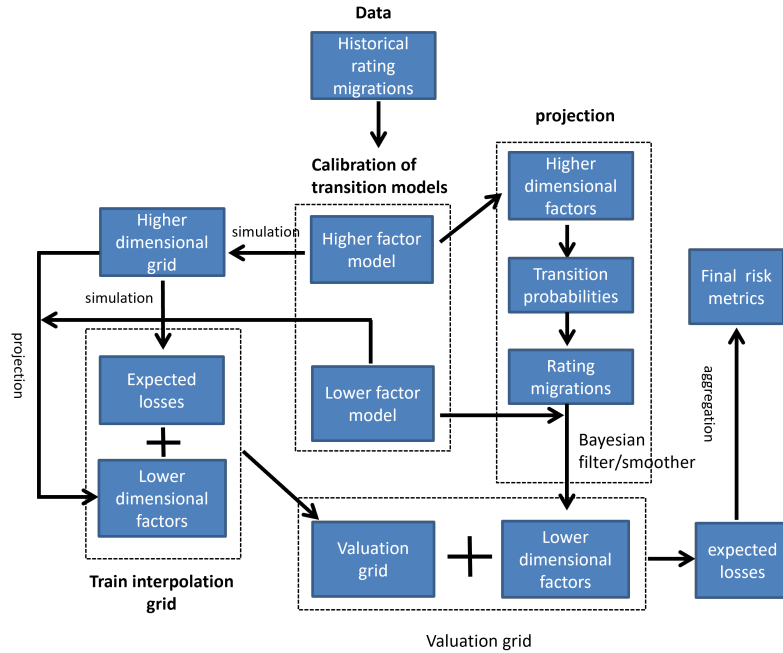


Figure 3.2: High level description of the algorithm

3.4.1 Credit Risk Economic Capital

Economic Capital, mainly for financial service forms, is a metric to determine the amount of capital a firm must hold to protect itself against economic risks (such as market risk, credit risk, and operational risk), given its mixture of assets and liabilities and its strategy to risk appetite. The risks are measured by the unexpected losses through client defaults and the asset value changes through the movements in the market parameters. Typically the losses are estimated based on a time horizon and a confidence level. The time horizon is usually one year and the confidence level is usually 99.9% or 99.95%, implying that the Economic Capital should cover the losses that would arise in case an adverse event would occur with a probability of 0.1% or 0.05% over one year horizon. In particular, the Credit Risk Economic Capital (CREC) is to measure the internal capital required to cover the unexpected credit losses under ECB ICAAP Guidelines, see European Central Bank (2018).

To measure the CREC, one particularly looks into two statistics: the expected loss (EL) and the 99.9% or 99.95% quantile of the portfolio loss. The CREC covers both default risk and migration risk. The default risk captures the losses from the default event within the time horizon, while the migration risk concerns the potential economic losses associated with non-default credit rating migrations. Although a non-default transition may not result in a direct loss under accrual accounting at the time horizon, it may result in an indirect loss in terms of the expected future cashflows, interest, principal and recoveries. According to this nature, the calculation of the CREC risk measures can be divided into two parts: the realised default losses within the time horizon and the potential losses after the time horizon. A Monte Carlo approach can be directly used to estimate the default losses within the time horizon, while a valuation grid is needed to compute the potential losses after the time horizon. Higher dimensional factors are desired for the simulation of the defaults and rating migrations up to

the horizon, due to the higher precision. Meanwhile, a dimensional factors are used as the inputs of the grids to compute the potential losses after horizon. Therefore, the proposed Bayesian dimension reduction approach can be implemented to project higher dimensional factors to lower dimensional factors for the valuation of the potential losses after the horizon.

3.4.2 Incremental Risk Charge

The Incremental Risk Charge (IRC), see European Banking Authority (2012), is a complement to the traditional stressed Value-at-Risk measured in Market Risk Economic Capital. It is an estimate of default and migration risk of unsecuritized credit products in the trading book. The concept and calculation of IRC is in general similar to CREC. The IRC also measures the default and migration risks at a 99.9% level of confidence over a one-year horizon. However, the CREC model mainly applies to the banking book, including exposures to government, institutions, corporates, retail exposures secured by real estate, and revolving retail exposures. It also covers the (one-year) counterparty credit risk of derivative instruments, securities financing transactions and hold-to-maturity assets in the trading book. By comparison, the IRC mainly concerns the credit risks of the bond and equity exposures in the trading book. Different from the booking book, in which the positions are held to maturity, the positions in the trading book are continuously bought and sold before their maturities. Therefore, unlike CREC, the IRC model differentiates the underlying traded instruments by liquidity horizon, with a minimum of three months⁶. A constant level of risk assumption⁷ is also imposed in IRC and ensures that all positions in the IRC portfolio are evaluated over the full one-year time horizon. Since the calculation of IRC also requires migration/default simulations for every liquidity until a one year capital horizon, and the migration risks need to be evaluated according to certain valuation function, the proposed Bayesian dimension reduction approach can be applied to make the calculations more efficient.

3.4.3 International Financial Reporting Standard 9

The International Financial Reporting Standard 9 (IFRS 9), see for example International Accounting Standards Board (2016), Board (2019), Gornjak (2020), is published by the International Accounting Standards Board (IASB). It addresses the accounting for financial instruments and it replaced the International Accounting Standard (IAS) 39 with regard to the methodology used to compute impairment provisions on financial instructions. IFRS 9 introduces a new impairment model based on the Expected Credit Loss (ECL), which is different from the Incurred Credit Loss (ICL) model under IAS 39. Under IFRS 9, institutions, as reporting entities, will have to recognize not only incurred credit losses but also losses that are expected in the future. In IFRS 9, two different ECLs are required, depending on the stages of impairment, see Bank for International Settlements (2017). In stage 1, the ECLs result from the default events that are possible within the next 12 months (12-month ECL). While in stage 2 or 3, lifetime ECLs are recognised. To obtain a 12-month ECL, a Monte Carlo simulation can be used based on the given PD and LGD models. While, for the lifetime

⁶Note that the capital horizon is still one year.

⁷The constant level of risk means the risk of the portfolio remains to be its initial risk level over the one-year horizon. In practice, This can be achieved by holding the constant portfolio over the liquidity horizon, rebalancing any default, downgraded, or upgraded positions at the beginning of each liquidity horizon, and rolling over any matured positions at the beginning of each horizon.

ECL, a grid approach is required⁸. Since the valuation of life time ECL usually involves a large number of factors from the PD and LGD models, the proposed Bayesian projection approach can be applied to reduce the dimensionality of the valuation grid for ECL and hence improve the efficiency of the computations.

3.4.4 Market risk

Market risk can be defined as the risk of losses in on and off-balance sheet positions arising from adverse movements in market prices. From a regulatory perspective, market risk stems from all the positions included in the trading book as well as from commodity and foreign exchange risk positions in the whole balance sheet. In market risk calculations, the VaR needs to be determined according to the future Profit-and-Loss (P&L) distribution of the portfolios. The value of the portfolio depends on the future values of the asset prices, equity market volatilities, interest rates, etc. Therefore, to obtain the P&L distribution, one has to simulate the future scenarios of (the risk factors of) the relevant asset prices, equity market volatilities or interest rates curves, and re-evaluate the portfolio values given the simulated scenarios. In order to efficiently re-evaluate the portfolio values, a valuation grid is usually used, see for instance Gibson and Pritsker (2000), Zamani et al. (2022). The valuation grid links the risk factors and the derivative values. However, many derivatives involve multiple underlyings, for instance the multi-asset options (Leentvaar and Oosterlee (2007), Ekedahl et al. (2007)). Therefore, the proposed Bayesian projection approach can be used in pricing such derivatives to avoid the “curse of dimensionality”.

3.4.5 Interest rate risk in the banking book

Interest rate risk in the banking book (IRRBB) refers to the current or prospective risk of the bank’s capital and earnings arising from adverse movements in interest rates that affect the bank’s banking book positions. When interest rates change, the present value and timing of future cash flows change. This in turn changes the underlying value of a bank’s assets, liabilities and off-balance sheet items and hence its economic value. Changes in interest rates also affect a bank’s earnings by altering interest rate-sensitive income and expenses, affecting its net interest income (NII). The calculation of the VaR for IRRBB is similar to the calculation of VaR for market risk. The P&L is generated by simulating the (risk factors of) interest rates curves and re-evaluating the balance sheet given the simulated scenarios. The re-evaluation of the balance sheet can be done by using the valuation grid method. In the past years since the credit and liquidity crisis of 2007, a multi-curve framework of interest rates was developed to replace the single-curve framework. The single-curve framework bootstrapped a single risk-free curve which represented both the cost of funding future cash flows and forward rates. This unique curve is used for both discounting and forecasting, while in the multi-curve framework, the discount curves and the forecasting curves are distinguished. As expected, the different forecasting tenor curve forecasts the future cashflows of different tenors and the discounting curve values these cashflows to today. More details regarding the single- and multi-curve framework can be found in, for instance, Henrard (2014), Pallavicini and Tarenghi (2010). The multi-curve framework introduces more curves in pricing the derivatives. For example, three curves are needed to price the Constant Maturity Swaps or the Basis Swaps,

⁸Sometimes the 12-month ECL is used to construct a proxy for the lifetime ECL in the simplified approach, see European Banking Authority (2021).

i.e. one discounting curve and two forecasting curves. If each curve requires a two-factor model, in total six factors are needed to construct a valuation grid for these derivatives. To decrease the dimensionality of the valuation grid, one can apply the proposed Bayesian projection approach.

4 Numerical studies

In the sequel we illustrate the performance of the projection approach by examples. We aim to forecast the bullets in Equation (2.4), and hence forecast the loss distribution of a loan. The chosen loan in the experiment is a unit bullet loan, of which the notional is 1. Consequently the principal repayment $PC = 0$ and $EAD_k = 1$ for $k = 1, \dots, n$. The coupon payments s_k are set to be equal to $s = \frac{1}{n}, k = 1, \dots, n$. For convenience and without loss of generality, discounting is not considered in this experiments. Hence Equation (2.4) simplifies to

$$L_1^{(r)} = \sum_{k=2}^n s \mathbb{E}[\text{PD}(k | \mathcal{F}_k) | \mathcal{F}_h] + \mathbb{E}[\text{PD}(n | \mathcal{F}_n) | \mathcal{F}_h] - \sum_{k=h+1}^n \mathbb{E}[(\text{PD}(k | \mathcal{F}_k) - \text{PD}(k-1 | \mathcal{F}_k))(1 - \text{LGD}_k) | \mathcal{F}_h]. \quad (4.1)$$

In all experiments we assume the maturity of the unit loan after a horizon of 30 periods, i.e. $n = 30$.

The value of the bullets depends on the transition and LGD models, therefore, we next introduce the transition and LGD models used in the experiments.

4.1 Transition model

The transition model we use belongs to the reduced form model. Let $k = 1, \dots, n$. Suppose we have $R > 2$ credit states. We denote by $T_{ij,k}$ the probability at time k of the transition from a credit state $i = 1, \dots, R$ to a credit state $j = 1, \dots, R$. We model the transition probabilities (or observed transition rates) using a logit function as follows,

$$T_{ij,k} = \frac{g_{ij} \exp(\theta_{ij,k})}{\sum_{j=1}^R g_{ij} \exp(\theta_{ij,k})}, \quad i, j = 1, \dots, R, \quad k = 1, \dots, n. \quad (4.2)$$

The $\theta_{ij,k} + \log g_{ij}$ are usually referred to as the log-odds for the transition from a credit state i to a credit state j , in which the g_{ij} are the parameters which indicate the level of the transition probabilities and $\theta_{ij,k}$ are the signals that describe the dynamics of the transition probabilities. Suppose, for clients in the i -th rating, that in $N_{i,k}$ independent trials the number of *observed* migrations of these clients to the j -th credit state is $m_{ij,k}$ for $j = 1, \dots, R$. It is easy to see that $m_{i,k} = [m_{i1,k}, \dots, m_{iR,k}]$ has a multinomial distribution, with log-density given by

$$\log p(m_{i,k} | T_k) = \sum_{j=1}^R m_{ij,k} \log T_{ij,k} + \log C_{i,k}, \quad (4.3)$$

where

$$C_{i,k} = \frac{N_{i,k}!}{\prod_{j=1}^R m_{ij,k}!}.$$

Plugging Equation (4.2) into Equation (4.3), one can rewrite the log-likelihood function (4.3) at time k as

$$\log p(m_{i,k}|\theta_{ij,k}) = \sum_{j=1}^R m_{ij,k}(\theta_{ij,k} + \log g_{ij}) - N_{i,k} \log \left(\sum_{j=1}^R g_{ij} \exp(\theta_{ij,k}) \right) + \log(C_k).$$

We assume that the migrations from a state i to a state j are driven by the *latent* common factors x (hence the migrations are also correlated). The latent process x is assumed to evolve linearly over time with a Gaussian error η , and therefore our transition model is given by:

$$\begin{aligned} m_{i,k} &\sim p(m_{i,k}|\theta_{ij,k}), & \theta_{ij,k} &= Kx_k, & k &= 1, \dots, n, \\ x_k &= Ax_{k-1} + \eta_k, & \eta_k &\sim \mathcal{N}(0, Q), \end{aligned} \tag{4.4}$$

where x_k is a $d \times 1$ random vector, K is a $(R-1) \times d$ -matrix, and A and Q are $d \times d$ -matrices. The $g = \{g_{ij}, i, j = 1, \dots, R\}$, K , A and Q are referred to as the parameters of the transition model. The initial distribution of x_k is assumed to be Gaussian with mean a_0 and (co)variance P_0 , i.e. $x_0 \sim \mathcal{N}(a_0, P_0)$.

This transition model belongs to the class of *state space models* of Section 3.1. The latent states x_t and the model parameters are usually estimated by the Bayesian filter introduced in Section 3.2. Specially, we estimate the latent states x_t by using the method of *mode estimation* (Durbin and Koopman (2012, Section 10.6)) based on the Kalman filter.

We suppose to have four different ratings with three performing ratings (P1, P2 and P3) and one default rating (D). We further assume that the default state is absorbing, which means that the probabilities of the transition from D to P1, P2 or P3 are zero. The set-up of the transition model in this example is the same as in Equations (4.4) and (4.2). In this experiment we have three models: one benchmark model and two testing models. The benchmark model is chosen to be a four factor model, i.e. the dimensionality of x_k is four. This model is treated as the ‘true’ model and used to generate the loss distribution which serves as a benchmark for comparison. The two testing models are chosen to be a two factor model and an one factor model. These two models are used to assess the performance of the proposed dimension reduction approach. All models are calibrated using the same migration data with a horizon of 120 periods. These migration data are simulated by using the benchmark model with parameters as follows. The parameter values are obtained from further classified real data, which we are not allowed to disclose.

$$\begin{aligned} A &= \text{diag}([0.6, 0.95, 0.9, 0.5]), \\ Q &= \text{diag}([0.6, 0.1, 0.1, 0.7]), \end{aligned}$$

$$\begin{aligned}
K &= \begin{pmatrix} k_{11,1} & k_{11,2} & k_{11,3} & k_{11,4} \\ k_{12,1} & k_{12,2} & k_{12,3} & k_{12,4} \\ k_{13,1} & k_{13,2} & k_{13,3} & k_{13,4} \\ k_{14,1} & k_{14,2} & k_{14,3} & k_{14,4} \\ k_{21,1} & k_{21,2} & k_{21,3} & k_{21,4} \\ k_{22,1} & k_{22,2} & k_{22,3} & k_{22,4} \\ k_{23,1} & k_{23,2} & k_{23,3} & k_{23,4} \\ k_{24,1} & k_{24,2} & k_{24,3} & k_{24,4} \\ k_{31,1} & k_{31,2} & k_{31,3} & k_{31,4} \\ k_{32,1} & k_{32,2} & k_{32,3} & k_{32,4} \\ k_{33,1} & k_{33,2} & k_{33,3} & k_{33,4} \\ k_{34,1} & k_{34,2} & k_{34,3} & k_{34,4} \\ k_{41,1} & k_{41,2} & k_{41,3} & k_{41,4} \\ k_{42,1} & k_{42,2} & k_{42,3} & k_{42,4} \\ k_{43,1} & k_{43,2} & k_{43,3} & k_{43,4} \\ k_{44,1} & k_{44,2} & k_{44,3} & k_{44,4} \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0.2 & 0.04 & 0.06 & 0.1 \\ 0.12 & -0.36 & 0.12 & -0.04 \\ 0.38 & -0.28 & -0.26 & -0.08 \\ -0.17 & 0.34 & 0.18 & -0.01 \\ 0 & 0 & 0 & 0 \\ 0.02 & -0.27 & 0.08 & -0.01 \\ -0.07 & -0.14 & 0.16 & -0.05 \\ -0.22 & -0.11 & 0.01 & 0.12 \\ -0.03 & -0.2 & 0.01 & 0.22 \\ 0 & 0 & 0 & 0 \\ -0.08 & 0.09 & -0.05 & -0.03 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \\
G &= \begin{pmatrix} g_{11} & g_{12} & g_{13} & g_{14} \\ g_{21} & g_{22} & g_{23} & g_{24} \\ g_{31} & g_{32} & g_{33} & g_{34} \\ g_{41} & g_{42} & g_{43} & g_{44} \end{pmatrix} = \begin{pmatrix} 0.95 & 0.03 & 0.0198 & 0.0002 \\ 0.05 & 0.9 & 0.04 & 0.01 \\ 0.05 & 0.12 & 0.78 & 0.05 \\ 0 & 0 & 0 & 1 \end{pmatrix}.
\end{aligned}$$

4.2 LGD model

We adopt ideas from Frontczak and Rostek (2015) and use the collateral to model the LGD as

$$\text{LGD}_k = \left(1 - \frac{c_k}{\text{EAD}_k}\right)^+, \quad (4.5)$$

where c_k is the value of the collateral at time k . At time k , the log-return of the collateral value is denoted LC_k , $\text{LC}_k = \frac{c_k}{c_{k-1}}$, $k \geq 1$. The log-return of the collateral is assumed to be modeled by an AR(1) process, see also Equation (A.2),

$$\text{LC}_k = 0.73 \text{LC}_{k-1} + z_k, \quad z_k \sim N(0, 0.04^2).$$

The value of the collateral, recall (4.5), determines the LGD as follows.

$$\begin{aligned}
\text{LGD}_k &= \left(1 - \frac{c_k}{\text{EAD}_k}\right)^+ \\
&= \left(1 - \frac{c_k}{c_0} * \frac{c_0}{\text{EAD}_0} * \frac{\text{EAD}_0}{\text{EAD}_k}\right)^+ \\
&= \left(1 - \frac{c_k}{c_0} * \frac{1}{\text{LTV}_0} * \frac{1}{n - k + 1}\right)^+.
\end{aligned}$$

Note that, in order to determine the LGD_k , the loan-to-value at the horizon LTV_0 should be defined. Moreover, for convenience, we further assume that the risk factors for the transition matrices and the log-returns of the collateral factors are independent, hence the default probabilities and the LGD are independent. Note that this assumption is not valid in real life practice. Actually, in practice the correlation between the PD and LGD is a very important

parameter in risk calculations. Under the independence assumption the loss in Equation (4.1) is simplified to

$$L_1^{(r)} = \sum_{k=2}^n s \mathbb{E} [\text{PD}(k | \mathcal{F}_k) | \mathcal{F}_h] + \mathbb{E} [\text{PD}(n | \mathcal{F}_n) | \mathcal{F}_h] - \sum_{k=2}^n \mathbb{E} [(\text{PD}(k | \mathcal{F}_k) - \text{PD}(k-1 | \mathcal{F}_k)) | \mathcal{F}_h] (1 - \mathbb{E} [\text{LGD}_k | \mathcal{F}_h]). \quad (4.6)$$

Given the assumed independence, we only need to approximate the multi-year expected probability of default $\mathbb{E}[\text{PD}_k^{(r)}]$ by Monte Carlo simulation, while the expected loss given default $\mathbb{E}[\text{LGD}_k]$ is analytically computed by using the Black-Scholes formula (A.6) of Corollary A.2.

In order to assess the performance of the proposed Bayesian projection approach, a *waterfall* terminology is used. In the waterfall, the analyses are done step by step as follows.

1. The Bayesian and PCA projection approach are compared when used to describe the transition probabilities. The two approaches project the four dimensional factors onto two dimensional factors respectively, and then the transition probabilities based on the two dimensional factors are compared with the original transition probabilities from the four dimensional factors. In particular, for the Bayesian projection approach, Algorithm 3.1 is applied.
2. The convergence analysis of the Monte Carlo simulation for expected PD is conducted to find out how many Monte Carlo scenarios are needed to obtain a reliable estimate of the multi-year expected PD.
3. Then the Bayesian and PCA projection approaches are used in the valuation grid to estimate the multi-year expected PD and accuracies are compared. In this case, Algorithm 3.2 is applied.
4. The closed-form Black-Scholes formula for the expected LGD is checked by comparing with the Monte Carlo estimation. The expected LGDs are used to construct the loss distribution according to Equation (4.1).
5. In the end, the proposed Bayesian approach, see Algorithm 3.2, is used to construct the loss distribution of a unit loan and the risk metrics (expected losses and value-at-risk) are estimated. The accuracy of the estimation is compared with the estimation from the PCA approach.

Table 1 summaries the waterfalls in the experiments Section 4.3.

4.3 Experiments

In this section we perform a number of numerical studies to assess and illustrate our projection approach.

Section	Description	Results
Section 4.3.1	Comparison of Bayesian and PCA projection approach on transition probabilities	Figure 4.1
Section 4.3.2	Convergence of Monte Carlo simulation for expected PD	Figure 4.2
Section 4.3.3	Comparison of Bayesian and PCA projection approach on multi-year PD estimation	Figures 4.3, 4.4
Section 4.3.4	Black-Scholes formula for expected LGD	Figure 4.5
Section 4.3.5	Comparison of Bayesian and PCA projection approach on loss distribution of the target unite bullet loan	Figures 4.6-4.9

Table 1: Waterfalls

4.3.1 Transition probabilities

Suppose the four dimensional model and the two dimensional model are calibrated on the data. In the experiment of this section, we assess the potential error introduced by the dimension reduction approaches, when projecting a higher dimensional transition model, in this case a four factor model, to a lower dimensional transition model, in this case, a two factor model. In particular, we apply the proposed projection approach, see Algorithm 3.1 and the PCA approach on the transition probability models and compare the performance of the two approaches. To achieve this, in total 100 different scenarios (risk factors and hence the transition probability matrices) are simulated using the four factor model. Then the proposed projection approach and the PCA approach are used to project the four dimensional factors onto two dimensional factors. For the proposed approach, the two factor model above is used and for the PCA approach the first two principal components are used to match with the number of factors. Based on the two dimensional factors, the transition probabilities can be recomputed according to the two factor model. These transition probabilities from the two factor model are compared with the original transition probabilities based on the four factor model and the relative difference between them are calculated. Figure 4.1 shows the relative difference (error) between the transition probabilities of the benchmark (four factor model) and the projections from the proposed approach and PCA approach. One can observe that the differences for the Bayesian projection approach are significantly smaller than those for the PCA approach.

4.3.2 Expected PD simulation

In order to create the training samples for the expected losses, the expected cumulative PD up to the maturity need to be approximated using Monte Carlo, see Equation (4.6). In the experiment of this section we investigate how many Monte Carlo scenarios are required to obtain a adequate approximation for the expected cumulative PD. For each rating $P1$, $P2$ and $P3$, we randomly create 1000 starting points for the simulation. These starting points can be treated as the points on the grids. Then starting at each point, the risk factor paths are simulated up to the maturity, in total 30 periods. The number of risk factor paths for each starting points varies from 500 to 20000. The expected cumulative PDs for each starting point are approximated by averaging over the risk factor paths. For the expected cumulative PD approximation 20000 paths are used as the benchmark, while the approximation using 500, 1000, 5000 and 10000 paths are used as the test sets. The relative difference of the

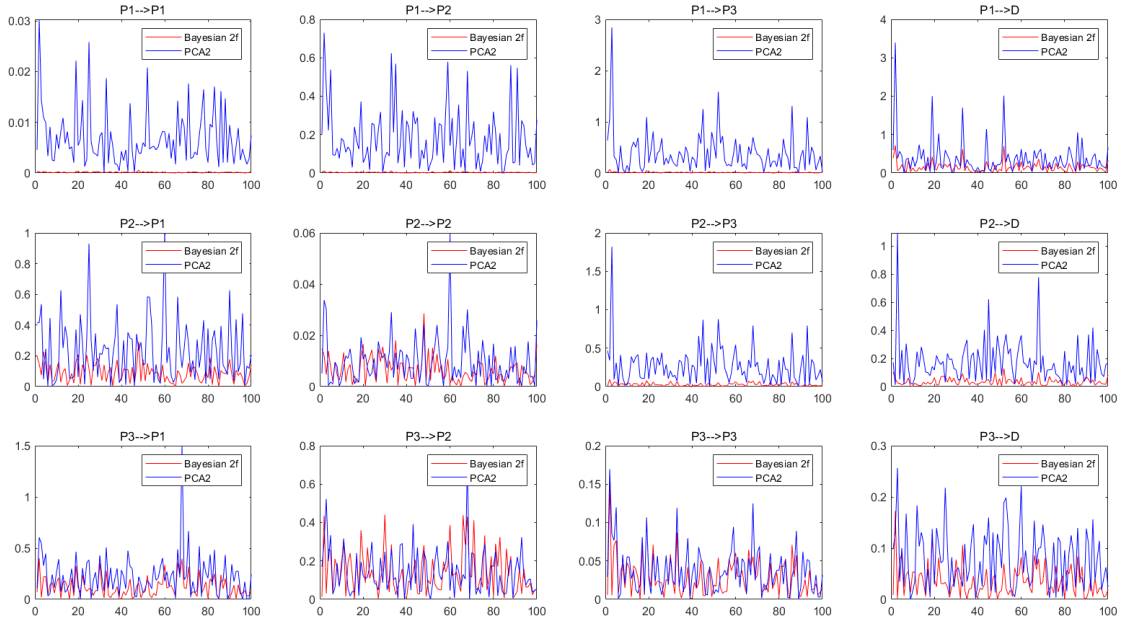


Figure 4.1: Relative difference of the transition probabilities: the red line is for the proposed Bayesian projection approach and the blue line is for the PCA. On the x -axis are the 100 simulated scenarios.

expected cumulative PD between the benchmark and the test sets are computed and the results are presented in Figure 4.2. The results indicate that the expected cumulative PD can be adequately approximated by using 1000 paths.

4.3.3 Experiments for the projection approach for expected PD

The purpose of the experiments is to assess the performance of the proposed Bayesian filter dimension reduction approach on valuating the losses, see Algorithm 3.2. Due to the independent assumption between the PD and LGD, the loss valuation is simplified such that the interpolation grid is only required for the expected PD valuation, see Equation (4.1). While the expected LGD can be easily computed by using the Black-Scholes formula in Appendix A. Therefore, in this section, we apply Algorithm 3.2 to compute the expected PD and the performance of the proposed Bayesian projection approach is assessed by comparing with the PCA approach. The assessment for the loss valuation will be presented later in Section 4.3.5.

Based on the four and two dimensional transition models, one can train the interpolation grid by simulating the expected PD for each point on the grid. As we described in Algorithm 3.2, the simulation of the expected PD uses the higher dimensional transition model while the grid points are in a lower dimensional space, obtained by a dimension reduction approach, i.e. the proposed Bayesian projection approach or the PCA approach. Then the valuation grid can be trained using the simulated expected PD and the projected factors. We compare the accuracy of the valuation grid based on the proposed Bayesian filter projection

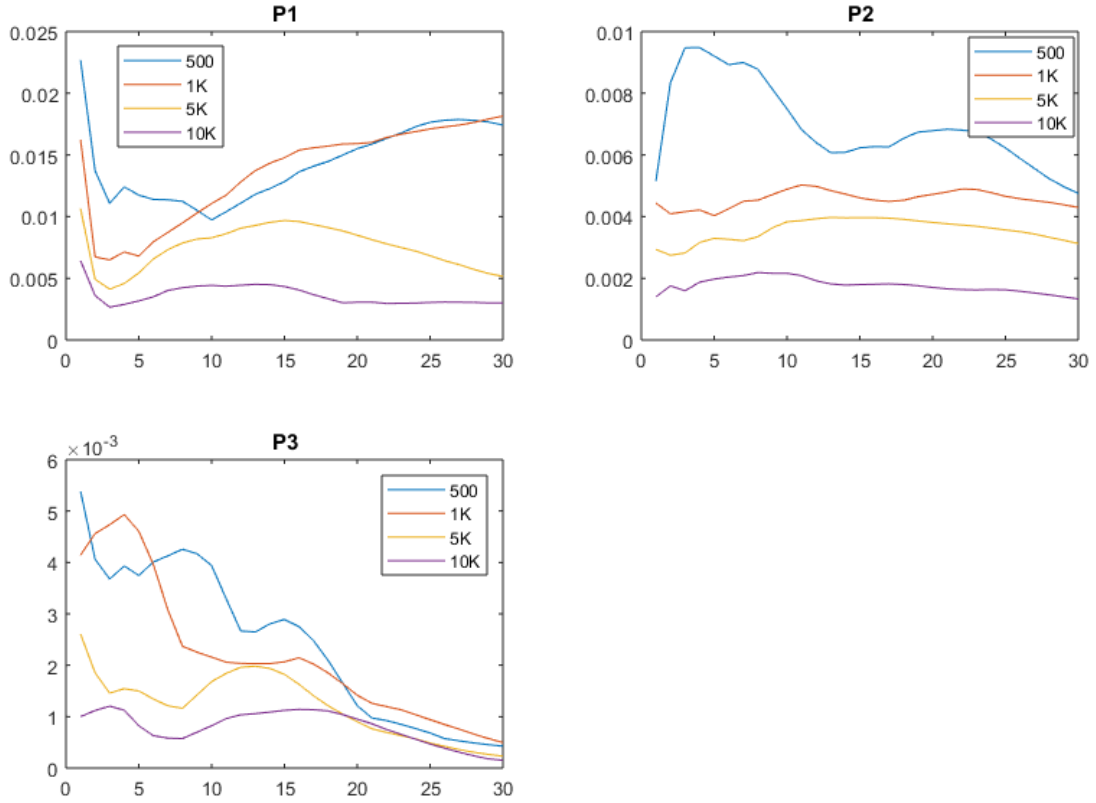


Figure 4.2: Maximum relative difference of the expected cumulative PD between 20000 MC paths and 500, 1000, 5000, 10000 paths: the y -axis is the maximum relative difference of the expected cumulative PD for the 1000 different starting points, on the x -axis are the time periods up to 30.

approach and the PCA approach. To build the valuation grid, the expected PD are simulated from the four factor model specified in Section 4.1.

The starting points of the simulations (i.e. the higher dimensional factors) consist of two parts: one contains the pre-determined points to cover the possible extreme simulated values of the factors, the other contains random simulated points to fill the density. In this experiments, the pre-determined points are chosen linearly spaced from minus four sigma to plus four sigma, i.e. for the i -th component of the factor, the linear spaced points are chosen from interval $[-4\sigma_i, 4\sigma_i]$, with σ_i the standard deviation of the i -th component of the factor. For each dimension, we chose 15 points. Moreover, 1000 more points are randomly simulated from the joint distribution of the four dimensional factors. Therefore, in total there are $15^4 + 1000 = 51625$ points in the valuation grid. Starting from the factors, each term structure of the expected PD is estimated by Monte Carlo simulation with 1000 paths. The term structure is considered from 1 to 30 periods. Meanwhile, the four dimensional factors are projected to lower dimensional factors using the Bayesian filter projection approach and the PCA approach. The chosen lower dimensional models are the one-factor and two-factor model. Once the valuation grids are trained, 1000 new four dimensional factors are simulated as the test set to assess the performance of different valuation grids based on different approaches. Based on these 1000 new four dimensional factors, the term structures of the expected PD

are estimated by using 20000 Monte Carlo paths. These Monte Carlo estimations are used as the benchmark for the comparison.

Figures 4.3 and 4.4 show the average (over the 20000 MC paths) relative difference on the term structure between the valuation grids and the Monte Carlo benchmarks. Specially, Figure 4.3 presents the comparison between the Bayesian filter projection using one factor model and the PCA projection using the first principal component, while Figure 4.4 presents the Bayesian filter projection of the two factor model and the PCA projection using the first two principal components. One observes that for one factor projection, in general PCA approach performs slightly better than the Bayesian filter projection approach. In particular, the Bayesian filter projection performs better for the first rating, but worse for rating 2 and rating 3. This is because the one factor model mainly focuses on the transitions of the first rating and hence the errors for the other two ratings are large. For the two factor projection, on the other hand, the Bayesian filter projection performs much better than the PCA approach. Comparing Figures 4.3 and 4.4, one observes very big improvements on the relative error for the Bayesian filter projection, while the improvement for the PCA projection is limited.

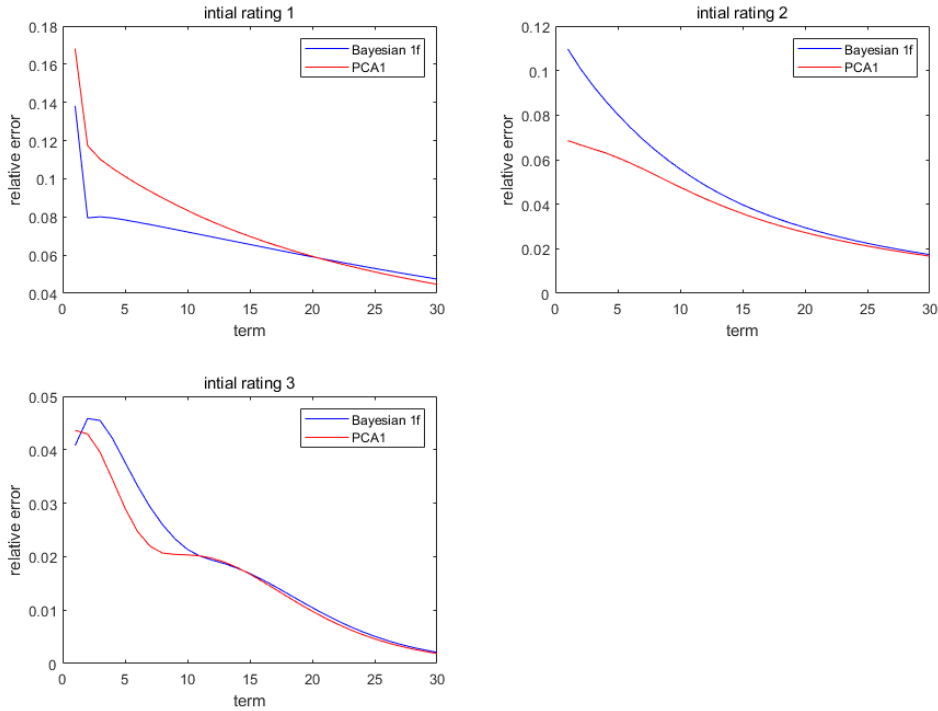


Figure 4.3: Average relative error on expected PD: Bayesian filter projection vs. PCA projection, one factor.

4.3.4 Expected LGD: Black-Scholes formula v.s. Monte Carlo

Before proceeding to the experiment on the loss valuation, in the experiment of this section, we illustrate the Black Scholes formula, the analytical solution of Equation (A.6) in Corollary A.2, for the expected LGD (ELGD), by comparing it to the Monte Carlo approximation. Note

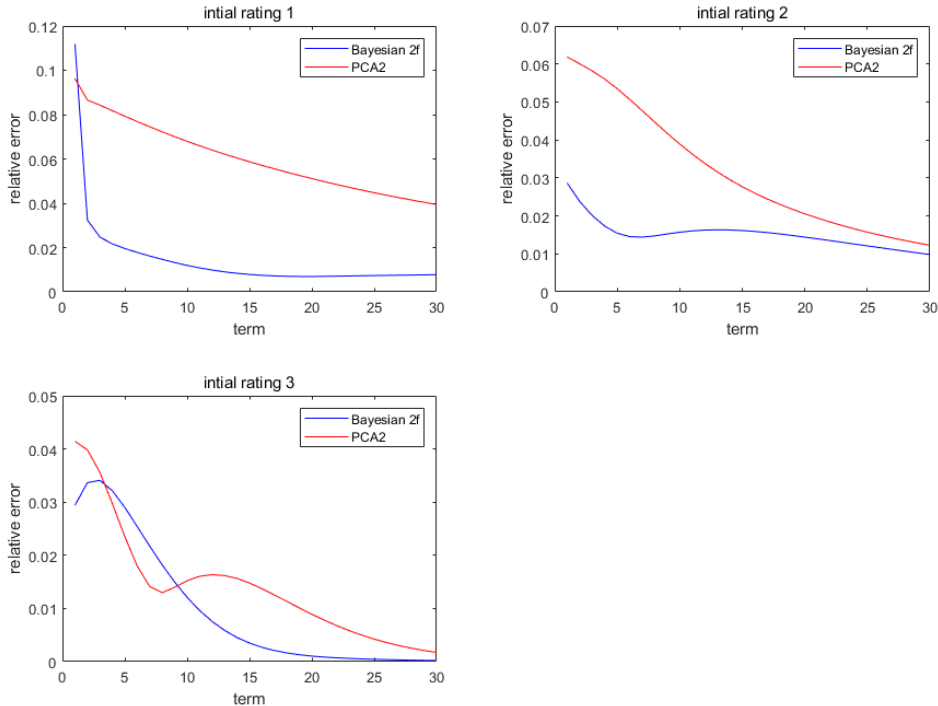


Figure 4.4: Average relative error on expected PD: Bayesian filter projection vs. PCA projection, two factors.

that the ELGD depends on the initial LTV and the initial log-return of the collateral. In this case, we choose the initial LTV to be 0.8, 1, 1.5 and 2, the initial log-return of the collateral to be -0.1, 0, and 0.1. Figure 4.5 shows the term structure of the ELGD computed by the Black-Scholes formula and the Monte Carlo simulation with 1 million scenarios. One can easily see that the outcomes of the two approaches align very well with each other.

4.3.5 Experiments for the projection approach for risk measures of the loss distribution

These experiments assess the accuracy of the proposed projection approach when estimating the quantiles of the loss distribution. In order to obtain a deeper insight about the performance, we investigate the loss distribution per initial rating, instead of the total loss distribution, which aggregates over different ratings. In this way we can compare the performance of the projection approaches for different initial ratings. To be more specific, the loss distribution per initial rating is constructed as follows.

1. Simulate one million scenarios (risk factors and collateral process) using the four factor model,
2. project the simulated risk factors into lower dimensional risk factors,
3. for each initial rating, given the lower dimensional factors, estimate the expected PD per scenario using the trained grid as in Section 4.3.3,

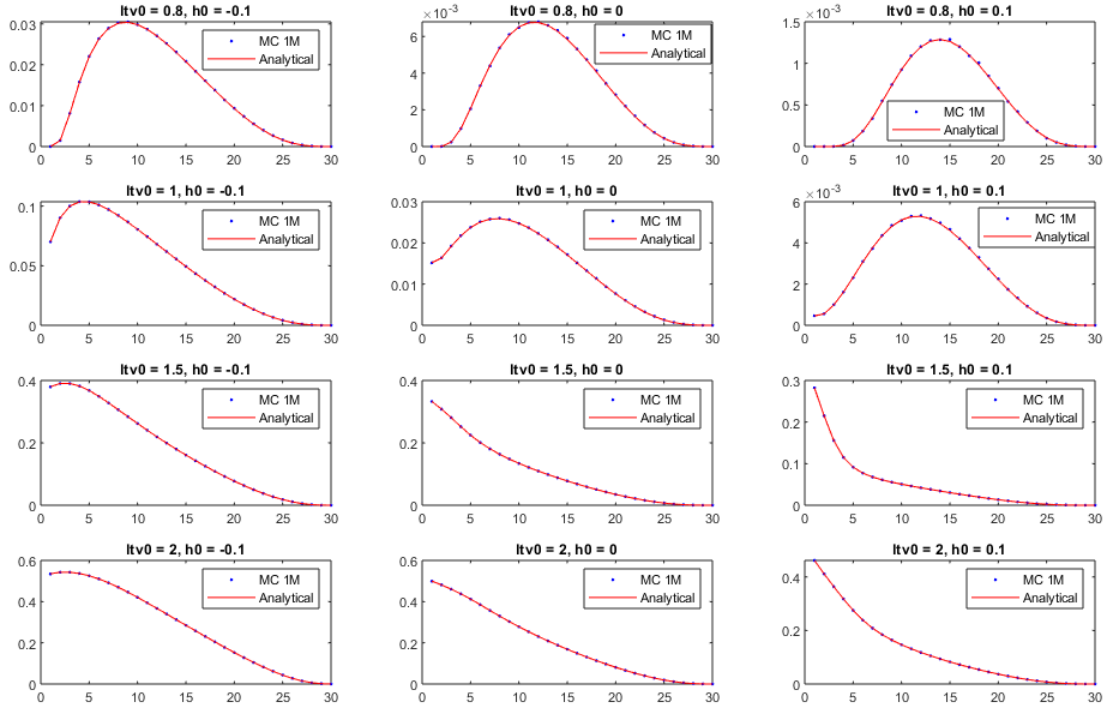


Figure 4.5: Term structure of ELGD: Black-Scholes (analytical) vs. Monte Carlo.

4. given the simulated collateral value, compute the expected LGD using the Black Scholes formula (A.6).
5. Note that the PD and LGD are assumed to be independent, therefore the loss per scenario can be evaluated, according to equation (2.2), by using the estimated expected PD and the expected LGD.

To illustrate the performance, the loss distribution based on the four factor model is used as a benchmark. In this benchmark, instead of using the projection approach and the interpolation grids, the expected PD per scenario in step 3 mentioned above is estimated directly by Monte Carlo simulation based on the four dimensional model itself. The chosen risk measures are expected loss, 95%-, 99%- and 99.9%-quantiles. We compute the absolute relative difference of these risk measures between the estimates from projection approaches and from the benchmark. Figures 4.6–4.9 present the comparison of the relative differences between different projection approaches for different risk measures: expected loss (EL), 95%-, 99%- and 99.9%- quantiles, respectively. In each figure, the dotted lines and the solid lines show the relative difference for the Bayesian projection approach and for PCA approaches, respectively, while the blue and red lines present the relative difference of one factor and two factor models (PCA), respectively.

In general, the results are in line with the results in Section 4.3.3. The Bayesian projection approach with the two factor model performs significantly better than the other candidates. Specially it has significantly smaller error than the PCA approach with the first two principal components. By contrast, the Bayesian projection approach with the one factor model does

not outperform the PCA approach with the first principal component. The reason is that the Bayesian projection approach with the one factor model focuses on the best rating, i.e. rating 1, which has the majority number of clients, while the PCA approach does not have this concentration. Therefore, one observes a better performance of one-factor Bayesian projection in the best rating but worse performance in rating 2 and 3. Another important observation is that, the Bayesian projection approach has a much better performance in the best rating than the PCA approach when estimating the more extreme quantile of the loss distribution. In particular one observes in Figure 4.9 that even the one factor Bayesian projection has comparable performance as the two-factor PCA projection in the best rating when estimating the 99.9%-quantile. This is a desired property in practice most of the exposures are located in the low-default-probability portfolios.

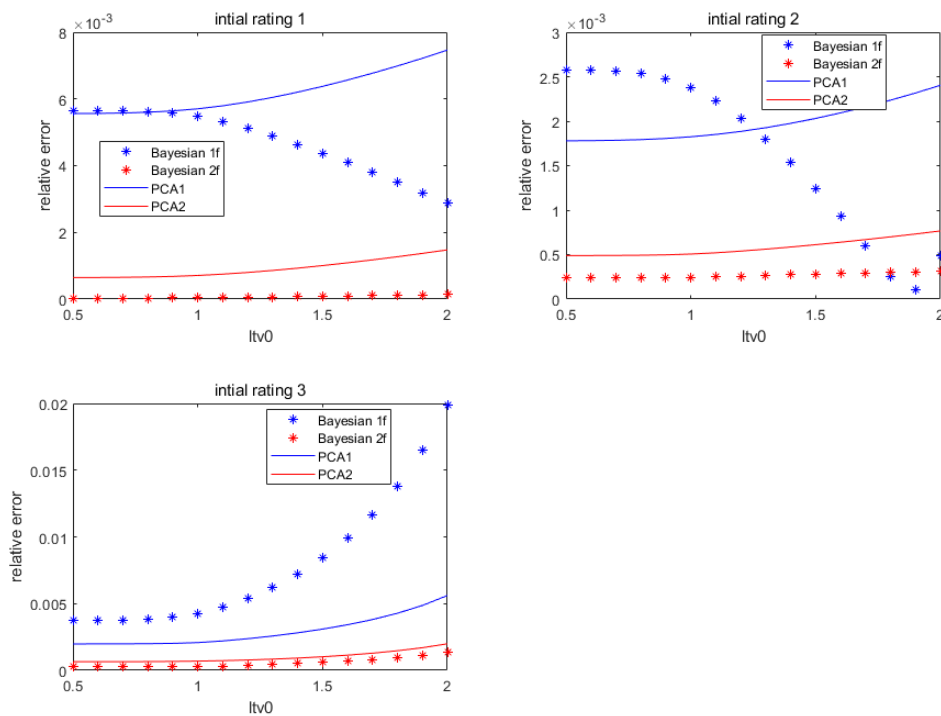


Figure 4.6: Absolute relative difference on expected loss: Bayesian filter projection vs. PCA projection.

5 Conclusions

In this paper we developed a dimension reduction methodology based on the Bayesian filter and smoother. This methodology is designed to achieve a fast and accurate loss valuation algorithm in the credit risk modelling, but it can also be extended to the valuation models of other risk types. The proposed methodology is generic, robust and can easily be implemented. Moreover, we showed the accuracy of the proposed methodology in the estimation of expected loss and value-at-risk by numerical experiments. The results suggest that, compared

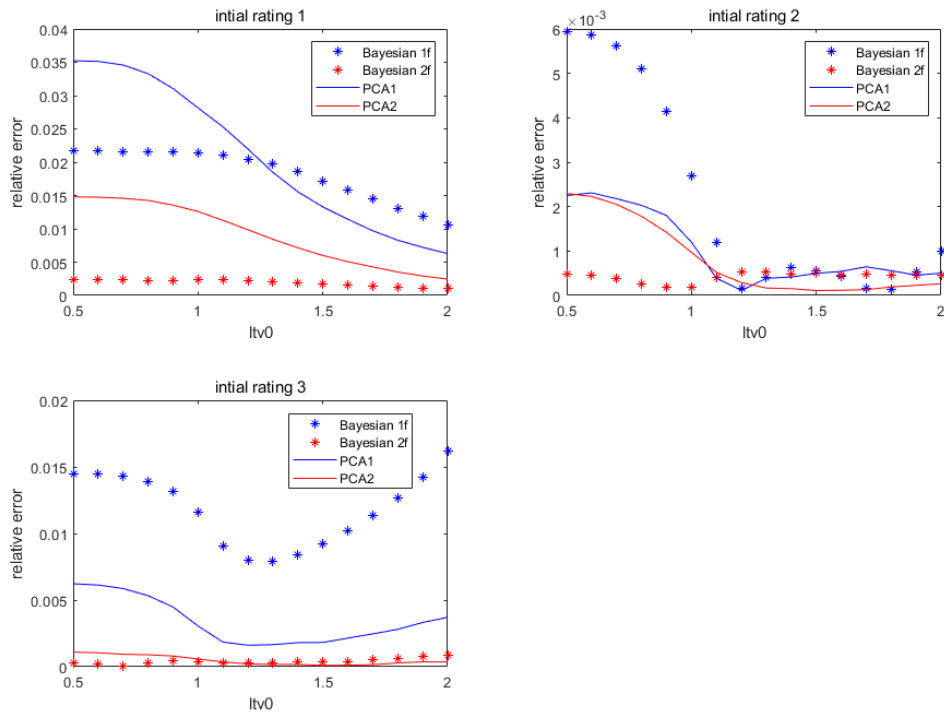


Figure 4.7: Absolute relative difference on 95%-quantile: Bayesian filter projection vs. PCA projection.

to the currently most used PCA approach, the proposed methodology provides more accurate estimation on the expected loss and value-at-risk of a loss distribution.

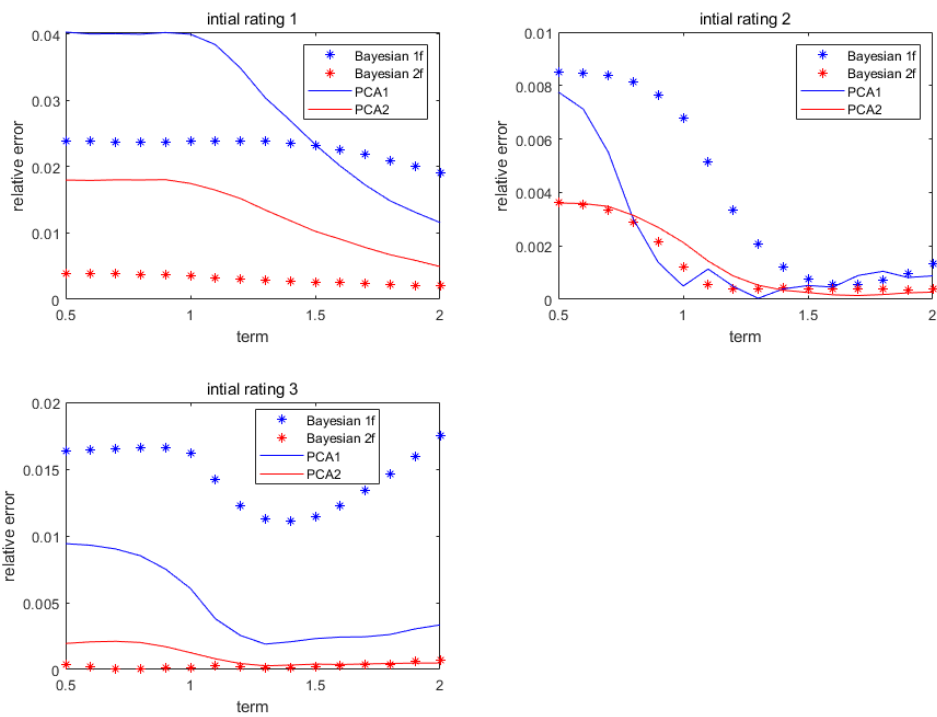


Figure 4.8: Absolute relative difference on 99%-quantile: Bayesian filter projection vs. PCA projection.

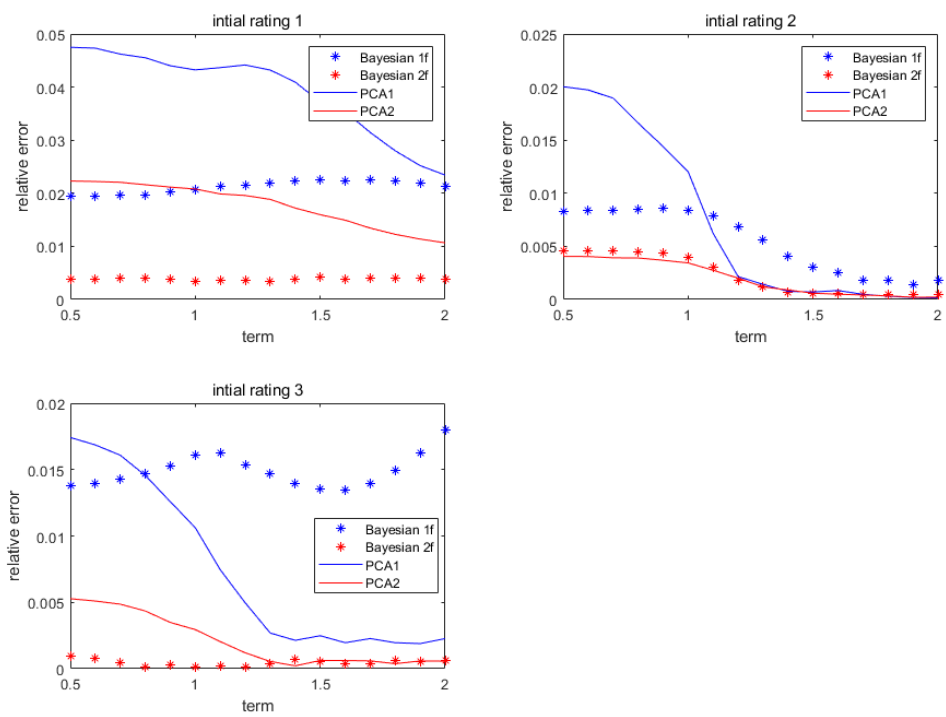


Figure 4.9: Absolute relative difference on 99.9%-quantile: Bayesian filter projection vs. PCA projection

A Analytic formula for expected loss-given-default

The LGD at time k conditional on the initial information, the sigma-algebra \mathcal{F}_0 , is defined by

$$\text{ELGD}_k = \mathbb{E}[\text{LGD}_k | \mathcal{F}_0].$$

In the present paper, we assume the LGD depends on the collateral value, see (4.5). Hence we have

$$\text{ELGD}_k = \mathbb{E} \left[\left(1 - \frac{c_k}{\text{EAD}_k}\right)^+ | \mathcal{F}_0 \right],$$

with c_k the value of collateral and EAD_k the exposure-at-default at time k . Since we assume the EAD_k is deterministic, the ELGD_k is determined by the stochastic nature of the collateral value c_k . The loan-to-value ratio is $\text{LTV}_0 = \frac{\text{EAD}_0}{c_0}$, hence we have

$$\begin{aligned} \text{ELGD}_k &= \mathbb{E} \left[\left(1 - \frac{c_k}{c_0} \frac{c_0}{\text{EAD}_0} \frac{\text{EAD}_0}{\text{EAD}_k}\right)^+ | \mathcal{F}_0 \right] \\ &= \mathbb{E} \left[\left(1 - \frac{c_k}{c_0} \frac{1}{\text{LTV}_0} \frac{\text{EAD}_0}{\text{EAD}_k}\right)^+ | \mathcal{F}_0 \right] \\ &= \mathbb{E} \left[\left(1 - K_k \frac{c_k}{c_0}\right)^+ | \mathcal{F}_0 \right], \end{aligned} \tag{A.1}$$

with $K_k = \frac{1}{\text{LTV}_0} \frac{\text{EAD}_0}{\text{EAD}_k}$. We suppose that the log-returns $\text{LC}_k = \log \frac{c_k}{c_{k-1}}$ of the collateral follow an AR(1) process,

$$\text{LC}_k = c + x\text{LC}_{k-1} + \sigma z_k, \quad z_k \sim N(0, 1), \quad k \geq 1, \tag{A.2}$$

with $|x| < 1$. Define the numbers a_k by the iteration

$$a_k = 1 + xa_{k-1}, \quad a_0 = 0, \tag{A.3}$$

which results in the analytic expression $a_k = \frac{1-x^k}{1-x}$. We now have the following result.

Proposition A.1. *For given LC_0 , it holds that*

$$\log \frac{c_t}{c_0} \sim N(\mu_t, \Omega_t), \tag{A.4}$$

with

$$\begin{aligned} \mu_t &= c \sum_{i=1}^t a_i + a_t x \text{LC}_0, \\ \Omega_t &= \sigma^2 \sum_{i=1}^t a_i^2. \end{aligned}$$

Proof. It follows from the recursion (A.2) that all LC_k are normal (conditional on LC_0). Hence for (A.4) to hold we only have to find mean and variance. We first prove a general result. Namely, for $s = 0, \dots, k$, the equation

$$\log \frac{c_k}{c_0} + \text{LC}_0 = c \sum_{i=1}^s a_i + a_{s+1} \text{LC}_{k-s} + \sum_{i=0}^{k-s-1} \text{LC}_i + \sigma \sum_{i=1}^s a_i z_{k-i+1} \tag{A.5}$$

holds. We now show (A.5) by induction w.r.t. the variable $s = 0, \dots, k$. When $s = 0$,

$$\log \frac{c_k}{c_0} = \sum_{i=1}^k \text{LC}_i,$$

which is Equation (A.5) for $s = 0$ by definition of the LC_i . Suppose then that Equation (A.5) holds for $s = t$. Using (A.2) with k replaced by $k - t$ and (A.3) with k replaced by $t + 2$, one gets starting from the induction assumption

$$\begin{aligned} \log \frac{c_k}{c_0} + \text{LC}_0 &= c \sum_{i=1}^t a_i + a_{t+1}(c + x\text{LC}_{k-t-1} + \sigma z_{k-t}) + \sum_{i=0}^{k-t-1} \text{LC}_i + \sigma \sum_{i=1}^t a_i z_{k-i+1} \\ &= c \sum_{i=1}^{t+1} a_i + a_{t+1}x\text{LC}_{k-t-1} + \text{LC}_{k-t-1} + \sum_{i=0}^{k-t-2} \text{LC}_i + \sigma \sum_{i=1}^{t+1} a_i z_{t-i+1} \\ &= c \sum_{i=1}^{t+1} a_i + a_{t+2}\text{LC}_{k-t-1} + \sum_{i=0}^{k-t-2} \text{LC}_i + \sigma \sum_{i=1}^{t+1} a_i z_{t-i+1}, \end{aligned}$$

which is indeed Equation (A.5) for $s = t + 1$, as was to be shown.

Taking $k = t$ and $s = k$ in Equation (A.5), we have, using (A.3) once more,

$$\begin{aligned} \log \frac{c_t}{c_0} &= -\text{LC}_0 + c \sum_{i=1}^t a_i + a_{t+1}\text{LC}_0 + \sigma \sum_{i=1}^t a_i z_{t-i+1} \\ &= c \sum_{i=1}^t a_i + xa_t\text{LC}_0 + \sigma \sum_{i=1}^t a_i z_{t-i+1}. \end{aligned}$$

Therefore we obtain the distribution of $\log \frac{c_t}{c_0}$, conditional on LC_0 ,

$$\log \frac{c_t}{c_0} \mid \text{LC}_0 \sim N(\mu_t, \Omega_t)$$

with $\mu_t = c \sum_{i=1}^t a_i + xa_t\text{LC}_0$ and $\Omega_t = \sigma^2 \sum_{i=1}^t a_i^2$, which finishes the proof. \square

As a corollary we obtain a Black-Scholes like formula for the expected loss given default.

Corollary A.2. *For the expected loss given default one has*

$$\text{ELGD}_t = \Phi(-d_2) - K_t e^{\mu_t + \frac{\Omega_t}{2}} \Phi(-d_1), \quad (\text{A.6})$$

with Φ the standard normal cumulative probability function, and

$$\begin{aligned} d_1 &= \frac{\log(K_t) + \mu_t + \Omega_t}{\sqrt{\Omega_t}}, \\ d_2 &= \frac{\log(K_t) + \mu_t}{\sqrt{\Omega_t}}. \end{aligned}$$

Proof. The proof is completely analogous to the proof of the classical Black-Scholes formula by realizing that Equation (A.1) gives $\text{ELGD}_t = \mathbb{E}[(1 - K_t \log Z)^+ \mid \mathcal{F}_0]$, where Z has the conditional normal distribution with mean μ_t and variance Ω_t as in Proposition A.1. \square

References

- C. Albanese, G. Campolieti, O. Chen, and A. Zavidonov. Credit barrier models. *Risk*, 16(6): 109–109, 2003.
- E.I. Altman. Default recovery rates and LGD in credit risk modelling and practice. In A. Lipton and A. Rennie, editors, *The Oxford Handbook of Credit Derivatives*, pages 39–65. Oxford Academic, 2011.
- E.I. Altman, B. Brady, A. Resti, and A. Sironi. The link between default and recovery rates: Theory, empirical evidence, and implications. *The Journal of Business*, 78(6):2203–2228, 2005.
- B.D.O. Anderson and J.B. Moore. *Optimal filtering*. Courier Corporation, 2012.
- E. Angelini, G. Di Tollo, and A. Roli. A neural network approach for credit risk evaluation. *The quarterly review of economics and finance*, 48(4):733–755, 2008.
- M.S. Arulampalam, S. Maskell, N. Gordon, and T. Clapp. A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking. *IEEE Transactions on signal processing*, 50(2):174–188, 2002.
- M. Avellaneda and J. Zhu. Distance to default. *Risk*, 14(12):125–129, 2001.
- Bank for International Settlements. IFRS 9 and expected loss provisioning – Executive Summary, 2017. URL <https://www.bis.org/fsi/fsisummaries/ifrs9.pdf>.
- Basle Committee. Credit risk modelling: Current practices and applications. *Basel Committee on Banking Supervision, Bank For International Settlements, Basel, Switzerland*, 1999. URL <https://www.bis.org/publ/bcbs49.htm>.
- Basle Committee. Basel II: International convergence of capital measurement and capital standards: A revised framework. *Basel Committee on Banking Supervision, Bank For International Settlements, Basel, Switzerland*, 2006. URL <https://www.bis.org/publ/bcbs128.htm>.
- R. Bellman. *Dynamic Programming*. Princeton University Press, 1957.
- C.M. Bishop. Neural networks and their applications. *Review of scientific instruments*, 65(6):1803–1832, 1994.
- ESRB-European Systemic Risk Board. Expected credit loss approaches in Europe and the United States: Differences from a financial stability perspective, 2019.
- M. Broadie and P. Glasserman. A stochastic mesh method for pricing high-dimensional american options. *Journal of Computational Finance*, 7:35–72, 2004.
- O. Cappé, S.J. Godsill, and E. Moulines. An overview of existing methods and recent advances in sequential Monte Carlo. *Proceedings of the IEEE*, 95(5):899–924, 2007.
- Z. Chen. Bayesian filtering: From Kalman filters to particle filters, and beyond. *Statistics*, 182(1):1–69, 2003.

- A. Chishti. Simulation of fixed-income portfolios using grids. *Algo Research Quarterly*, 2(2): 41–50, 1999.
- N. Chopin, P.E. Jacob, and O. Papaspiliopoulos. SMC²: an efficient algorithm for sequential analysis of state space models. *Journal of the Royal Statistical Society. Series B. (Methodological)*, 75(3):397–426, 2013. URL <https://doi.org/10.1111/j.1467-9868.2012.01046.x>.
- C.K. Chui and G. Chen. *Kalman filtering with real-time applications*. Springer, Cham, Fifth edition, 2017. URL <https://doi.org/10.1007/978-3-319-47612-4>.
- D. Crisan and J. Míguez. Nested particle filters for online parameter estimation in discrete-time state-space Markov models. *Bernoulli*, 24(4A):3039–3086, 2018. URL <https://doi.org/10.3150/17-BEJ954>.
- G. Delianedis and R.L. Geske. Credit risk and risk neutral default probabilities: information about rating migrations and defaults. *Available at SSRN 424301*, 2003.
- A. Doucet and A. M. Johansen. A tutorial on particle filtering and smoothing: fifteen years later. In *The Oxford handbook of nonlinear filtering*, pages 656–704. Oxford Univ. Press, Oxford, 2011.
- D. Duffie and K.J. Singleton. *Credit risk: pricing, measurement, and management*. Princeton University Press, 2004.
- J. Durbin and S.J. Koopman. *Time series analysis by state space methods*. Oxford University Press, 2012.
- G.A. Einicke and L.B. White. Robust extended Kalman filtering. *IEEE Transactions on Signal Processing*, 47(9):2596–2599, 1999.
- E. Ekedahl, E. Hansander, and E. Lehto. Dimension reduction for the Black-Scholes equation. *Department of Information Technology, Uppsala University*, 2007.
- European Banking Authority. EBA guidelines on the incremental default and migration risk charge (IRC), 2012. URL <https://www.eba.europa.eu/sites/default/documents/files/documents/10180/104828/938f9f0d-ec75-4874-9835-870e7cfda648/EBA-BS-2012-79--GL-on-IRC-.pdf?retry=1>.
- European Banking Authority. IFRS 9 implementation by EU institutions - monitoring report, 2021. URL https://www.eba.europa.eu/sites/default/documents/files/document_library/Publications/Reports/2021/1024609/IFRS9%20monitoring%20report.pdf.
- European Central Bank. ECB guide to the internal capital adequacy assessment process (ICAAP), 2018. URL https://www.bankingsupervision.europa.eu/ecb/pub/pdf/ssm.icaap_guide_201811.en.pdf.
- T. Evans and P. Nair. Scalable Gaussian processes with grid-structured eigenfunctions (gp-grief). In *International Conference on Machine Learning*, pages 1417–1426. PMLR, 2018.

- R. Frey and A.J. McNeil. Dependence modelling, model risk and model calibration in models of portfolio credit risk. *Preprint*, 2002.
- R. Frontczak and S. Rostek. Modeling loss given default with stochastic collateral. *Economic Modelling*, 44:162–170, 2015.
- J. Frye. Collateral damage: A source of systematic credit risk. *Risk*, 13(4):91–94, 2000a.
- J. Frye. Depressing recoveries. *Risk*, 13(11):108–111, 2000b.
- M.S. Gibson and M. Pritsker. Improving grid-based methods for estimating value at risk of fixed-income portfolios. Technical Report 25, Board of Governors of the Federal Reserve System, Finance and Economics Working Paper, 2000.
- M. Gornjak. *Literature Review of IFRS 9 and Its Key Parameters*, pages 463–477. Expanding Horizons: Business, Management and Technology for Better Society. ToKnowPress, 2020. URL <https://ideas.repec.org/h/tkp/mk1p20/463-477.html>.
- V. Gumbo. EAD Modeling for credit card. *Pioneer Journal of Theoretical and Applied Statistics*, 6(1):43–51, 2013. URL http://www.pspchv.com/content_1_PJTAS-vol-6-issues-1.html.
- J. He, A. Khedher, and P. Spreij. A Kalman particle filter for online parameter estimation with applications to affine models. *Statistical Inference for Stochastic Processes*, 24:353–403, 2021.
- M. Henrard. *Interest rate modelling in the multi-curve framework: Foundations, evolution and implementation*. Springer, 2014.
- M. Hillebrand. Modelling and estimating dependent loss given default. *Risk*, 19(9):120, 2006.
- B. Horvath, A. Muguruza, and M. Tomas. Deep learning volatility: a deep neural network perspective on pricing and calibration in (rough) volatility models. *Quantitative Finance*, 21(1):11–27, 2021.
- J.C. Hull and A.D. White. Valuing credit default swaps II: Modeling default correlations. *The Journal of derivatives*, 8(3):12–21, 2001.
- International Accounting Standards Board. IFRS 9 financial instruments, 2016. URL <http://eifrs.ifrs.org/eifrs/bnstandards/en/2016/ifrs09.pdf>.
- F. Jamshidian and Y. Zhu. Scenario simulation: Theory and methodology. *Finance and stochastics*, 1(1):43–67, 1996.
- R.A. Jarrow and S.M. Turnbull. Pricing derivatives on financial securities subject to credit risk. *The journal of finance*, 50(1):53–85, 1995.
- M. Jeanblanc and Y. Lecom. Reduced form modelling for credit risk. *Available at SSRN 1021545*, 2008.
- M. Jeanblanc and L. Li. Characteristics and constructions of default times. *SIAM Journal on Financial Mathematics*, 11(3):720–749, 2020.

- M. Jeanblanc and M. Rutkowski. Models for default risk: an overview. *Mathematical finance: theory and practise*, J. Yong and R. Cont, Eds. Higher education press, 1999.
- I.T. Jolliffe. Principal component analysis: a beginner's guide - I. Introduction and application. *Weather*, 45(10):375–382, 1990.
- S.J. Koopman, A. Lucas, and A. Monteiro. The multi-state latent factor intensity model for credit rating transitions. *Journal of Econometrics*, 142(1):399–424, 2008.
- C.C.W. Leentvaar and C. W. Oosterlee. Multi-asset option pricing using a parallel fourier-based technique. *Reports of the Department of Applied Mathematical Analysis*, 07-12, 2007.
- R.C. Merton. On the pricing of corporate debt: The risk structure of interest rates. *The Journal of Finance*, 29(2):449–470, 1974.
- P. Nickell, W. Perraudin, and S. Varotto. Stability of rating transitions. *Journal of Banking & Finance*, 24(1-2):203–227, 2000.
- P. Pagnotoni. Neural network models for bitcoin option pricing. *Frontiers in Artificial Intelligence*, 2:5, 2019.
- A. Pallavicini and M. Tarenghi. Interest-rate modeling with multiple yield curves. *arXiv:1006.4767*, 2010.
- S.J. Press. *Subjective and objective Bayesian statistics*. Wiley Series in Probability and Statistics. Wiley-Interscience [John Wiley & Sons], Hoboken, NJ, Second edition, 2003. ISBN 0-471-34843-0. Principles, models, and applications, With contributions by Siddhartha Chib, Merlise Clyde, George Woodworth and Alan Zaslavsky.
- C.E. Rasmussen and C.K.I. Williams. *Gaussian processes for machine learning*. Adaptive Computation and Machine Learning. MIT Press, Cambridge, MA, 2006. ISBN 978-0-262-18253-9.
- C.P. Robert. *The Bayesian choice, From decision-theoretic foundations to computational implementation*. Springer Texts in Statistics. Springer, New York, second edition, 2007. ISBN 978-0-387-71598-8.
- S. Särkkä. *Bayesian filtering and smoothing*, volume 3 of *Institute of Mathematical Statistics Textbooks*. Cambridge University Press, Cambridge, 2013. ISBN 978-1-107-61928-9. doi: 10.1017/CBO9781139344203. URL <https://doi.org/10.1017/CB09781139344203>.
- J.F. Thompson, B.K. Soni, and N.P. Weatherill. *Handbook of grid generation*. CRC press, 1998.
- E.N.C. Tong, C. Mues, I. Brown, and L.C. Thomas. Exposure at default models with and without the credit conversion factor. *European Journal of Operational Research*, 252(3): 910–920, 2016.
- J.H. van Schuppen. Stochastic realization problems. In *Three decades of mathematical system theory*, volume 135 of *Lect. Notes Control Inf. Sci.*, pages 480–523. Springer, Berlin, 1989. doi: 10.1007/BFb0008474. URL <https://doi.org/10.1007/BFb0008474>.

- E.A. Wan and A.T. Nelson. *Dual Extended Kalman Filter Methods*, chapter 5, pages 123–173. John Wiley & Sons, Ltd, 2002. ISBN 9780471221548. doi: 10.1002/0471221546.ch5. URL <https://onlinelibrary.wiley.com/doi/abs/10.1002/0471221546.ch5>.
- E.A. Wan and R. van der Merwe. *The Unscented Kalman Filter*, chapter 7, pages 221–280. John Wiley & Sons, Ltd, 2002. ISBN 9780471221548. doi: 10.1002/0471221546.ch7. URL <https://onlinelibrary.wiley.com/doi/abs/10.1002/0471221546.ch7>.
- J.Z. Wei. A multi-factor, credit migration model for sovereign and corporate debts. *Journal of International Money and Finance*, 22(5):709–735, 2003.
- G. Welch and G. Bishop. An introduction to the Kalman filter. Technical report, University of North Carolina at Chapel Hill, 1995.
- J. Wendin and A.J. McNeil. Dependent credit migrations. *Journal of credit risk*, 2(3):87–114, 2006.
- S. Zamani, A. Chaghazardi, and H. Arian. Pathwise grid valuation of fixed-income portfolios with applications to risk management. *Heliyon*, 8(7):e09880, 2022.