Numerical methods for structural credit models with mutual liabilities

Vadim Kaushansky

Mansfield College

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy

Hilary 2019

Acknowledgements

I would like to express my sincere gratitude to my supervisors, Prof. Alexander Lipton and Prof. Christoph Reisinger, for their guidance, patience, encouragement, sharing their knowledge, and building me as a researcher through my doctoral studies.

I would like to thank my co-supervisor Dr. Marek Musiela for organizing my DPhil, sharing his knowledge of quantitative finance, and interesting discussions.

I would like to thank Prof. Sam Howison, Dr. Athena Picarelli, Prof. Ben Hambly, Dr. Jeff Dewynne for their valuable feedback from my Transfer and Confirmation of Status.

I would also thank my colleagues, Zhenru Wang, Yixuan Wang, Wei Fang, Yufei Zhang, Andrei Cozma, and Matthieu Mariapragassam, with whom I got a chance to share my DPhil experience. A special thank to my colleague, Andreas Sojmark, for interesting and fruitful discussions on several research topics.

I am grateful to my friends, Maxim, Vitaly, Vadim, and Filip for their help and support during all the difficult times.

Last, but most importantly, I would like to thank my family. I am infinitely grateful to my parents, Tatyana and Yakov, my sister and brother, Inna and Sasha, and my girlfriend Anna, who consistently supported and encouraged me and without whom this thesis would not be written.

Abstract

In this thesis, we study a structural default model of an interconnected banking system with mutual liabilities. We develop novel numerical and analytical methods for the computation of important characteristics of the model and assess their impact on counterparty credit risk. First, we study the case when the banks' asset values follow a pure diffusion process and develop analytical and semi-analytical methods for default probabilities, which are the first of their kind in the three-dimensional case. Next, we consider a jump-diffusion model and develop new splitting finite difference methods for the resulting partial integro-differential equations, which we test numerically in the case of two banks. Due to the curse of dimensionality, these methods are not extendable practically to a large number of banks. Therefore, we consider a mean-field approximation of the model in this case, which leads to a McKean–Vlasov equation where the drift term depends on the absorption rate on the boundary. We develop and analyse a novel simulation method and an alternative approach based on reduction of the corresponding nonlinear, nonlocal PDE to a coupled system of Volterra integral equations.

Contents

List of Figures

Chapter 1

Introduction

Since the global financial crisis around 2008, the estimation and mitigation of counterparty credit risk has become a pillar of financial risk management. The impact of such risks on financial derivatives is explicitly acknowledged by a valuation adjustment. Moreover, a particular concern to regulators and central banks is the impact of default of an entity on the financial system – credit contagion, which might lead to a chain of defaults of the whole banking system. The possibility of events that might cause instability of the banking system is called *systemic risk*, the risk of collapse of an entire financial system.

Different approaches are used to model systemic risk. One approach is based on network analysis (Eisenberg and Noe (2001); Gai and Kapadia (2010); Glasserman and Young (2016); Allen and Gale (2000); Cont et al. (2010); Gai et al. (2011)). We call this a static approach since it models the current behaviour of the banking network based on interbank data. Barker et al. (2016) and Lipton (2017) (Chapter 16) measured systemic risk in Central Clearing House Networks.

Another approach is dynamic and based on credit default models. There are two main approaches for this: the *reduced form approach* and the *structural approach*. We are going to briefly review both approaches. Then, we briefly discuss *mean-field* approximations in the structural framework for a large number of banks.

The focus of this thesis is to develop and analyse semi-analytical and numerical methods for a structural default model with contagion through mutual liabilities for a small number of banks and in the large-pool limit.

1.1 Literature overview

Reduced form models The idea of reduced-form models is not to go deep into the economic reasons a default event, but to assume that the default comes purely randomly and to model the default rate, given that the default has not occurred yet, called the hazard function. There are many papers devoted to the reduced-form approach. Some of the ones mainly used in practice are Artzner and Delbaen (1995), Jarrow and Turnbull (1995), Duffie and Singleton (1997), Duffie and Singleton (1999), Lando (1998). A full review of this approach can be found in Bielecki and Rutkowski (2013).

The main advantage of the reduced-form approach is that it can be quite well calibrated to market quotes and, therefore, used for pricing credit derivatives. But, on the

other hand, the approach does not have any economic interpretation of the cause of the default event.

This approach was also extended by Spiliopoulos (2015) and Giesecke et al. (2013), who considered a large-pool limit of credit portfolios and measured systemic risk using this approach.

Structural models The second class of models are structural models. In contrast to reduced-form models, the structural framework deduces default from the balance sheet (i.e., from the information about assets and liabilities). We further follow Lipton and Sepp (2013) for the review of structural models.

Originally, the structural framework was developed in Merton (1974) by what is now known as the Merton model. There, it is assumed that firms' asset values are driven by Geometric Brownian Motion and the firm borrows a zero-coupon bond with face value F and time to maturity T. The firm defaults at time T if its asset value at maturity time T is less than the corresponding face value. Then, the corresponding bond price at time t or default probability can be computed via the standard Black and Scholes framework.

Many authors proposed various extensions of this basic model. One of them was introduced in Black and Cox (1976), where the authors introduced continuous default monitoring. Thereby, the default probability can be computed using pricing formulas for barrier options in the Black and Scholes framework. Also, a more complicated debt structure was considered, for instance in Kim et al. (1993), Nielsen et al. (1993), Leland (1994), Longstaff and Schwartz (1995), Leland and Toft (1996), Albanese and Chen (2005). Most recent research papers concentrated on extending the model in order to be able to generate the high short-term CDS spreads typically observed in the market. This has been done in different ways: Hyer et al. (1999), Hull and White (2001), Avellaneda and Zhu (2001) made default barriers curvilinear, Finger et al. (2002) made default barriers stochastic, and Zhou (2001b), Hilberink and Rogers (2002), Lipton (2002), Lipton et al. (2007), Sepp (2004), Sepp (2006), Cariboni and Schoutens (2007), Feng and Linetsky (2008) incorporated jumps into the firm values' dynamics.

Several authors also considered multi-dimensional extensions of the structural model: Zhou (2001a), Hull and White (2001), Haworth (2006), Haworth et al. (2008), Valuzis (2008) considered bivariate correlated log-normal dynamics for two firms; Li (2000) considered the Gaussian copula description of correlated default times in multi-dimensional structural models; Kiesel and Scherer (2007) studied a multi-dimensional structural model and proposed a mixture of semi-analytical and Monte–Carlo methods for pricing and calibration. Lipton and Sepp (2013) built a general multi-dimensional structural model aimed in quantitative estimation of counterparty risk. Ballotta et al. (2019) considered another type of structural model, where they model the assets as general correlated Lévy process.

Recent studies (see e.g. David and Lehar (2014)) showed that the banking system is highly interconnected and large mutual obligations can have a severe effect on the stability of the system. Itkin and Lipton (2017), Itkin and Lipton (2015), and Lipton (2016) extended the Lipton–Sepp framework for the case of mutual obligations. Of the various channels of such systemic risk described in Hurd (2016), we focus here on dependencies

through asset correlation and interbank liabilities. Specifically, we consider the extended structural default model introduced in Lipton (2016), where asset values are assumed to follow stochastic processes with correlated diffusion and jump components, and where mutual liabilities can lead to default cascades.

Itkin and Lipton (2017) considered the model without jumps and obtained explicit expressions for several quantities of interest including the joint and marginal survival probabilities as well as CDS and FTD prices. They demonstrate that mutual liabilities can have a large impact on the survival probabilities of banks. Thus, a shock of one bank can cause ripples through the whole banking system. On the other hand, Itkin and Lipton (2015) developed a finite difference method for the resulting partial integro-differential equation (PIDE) where the integral term results from a fairly general correlated Lévy process in the jump component.

Mean–field models When the banking system is large and homogenous, and only macroscopic quantities are of interest, one can consider a large-pool approximation of the banks' asset value processes (technically, by taking the limit of their empirical measure for an infinite number of banks).

The approach based on mean-field limit is commonly used in various areas of applied mathematics. Full details with various examples are given in Carmona and Delarue (2018).

In banking system modeling context, it was first studied by Bush et al. (2011). Following on, Nadtochiy and Shkolnikov (2017) and Hambly et al. (2018) took into account interaction effects by considering a particle system with positive feedback from the firms' defaults. This leads to McKean–Vlasov type equations, which model a typical representative of the banking system whose dynamics depends on the losses in the wider system.

Hambly et al. (2018) assumed zero correlation between banks with linear dependence of the interaction on the loss function and obtained existence and uniqueness of the solution, with the necessity of blow-ups for large enough interaction parameter. Later, Hambly and Sojmark (2018) and Ledger and Sojmark (2018) introduced positive correlation between banks, while Nadtochiy and Shkolnikov (2017) and Nadtochiy and Shkolnikov (2018) considered a nonlinear dependence through the loss function. Very recently, Ichiba et al. (2018) derived a McKean-Vlasov SDE in nonlinear jump-diffusion form for the average bank reserves in an interacting banking system with local and mean-field default intensities.

In the next chapter, we show the connection between the structural model framework of Lipton (2016) and the mean-field models described above. We show that by taking the number of banks to infinity, this leads to the same McKean–Vlasov equation.

1.2 Main contributions

The main contributions of the thesis are the following:

- We find a semi-analytical formula for the three-dimensional transition probability of Brownian motion in the positive octant with absorbing boundaries. The density is hereby represented as a series of special functions containing unknown parameters, which can be found as the solution of a nonlinear eigenvalue problem. We apply the results to the default model for three banks without jumps, and obtain expressions for joint survival probability, CVA and DVA for a CDS. The results have been published in Kaushansky et al. (2018b) and are given in Sections 3.6–3.8.

- We develop a finite-difference method, an extension of the Hundsdorfer-Verwer scheme, for a partial integro-differential equation (PIDE) resulting from a jump-diffusion model, study its stability and consistency. The method gives second order convergence in both time and space variables and is unconditionally stable. By applying the finite-difference method, we compute various model characteristics, such as joint and marginal survival probabilities, CDS and FTD spreads, as well as CVA and DVA, and estimate the impact of jumps on the results. For a more sophisticated analysis, we calibrate the model to the market, and demonstrate a sizeable impact of jumps on joint and marginal survival probabilities in the case of two banks. The development of numerical methods which are feasible for larger systems of banks appears to be an important future research direction. The results have been published in Kaushansky et al. (2018a) and are given in Sections 4.2–4.3.

- We develop particle methods with explicit timestepping for the simulation of the McKean–Vlasov equation arising in the infinite-dimensional model. Convergence with a rate up to 1 in the timestep is shown under a condition on the model parameters and time horizon, when the loss function is differentiable. Experimentally, the method also converges in the blow-up regime, and the variance of estimators is inversely proportional to the number of samples. The results have been published in Kaushansky and Reisinger (2018) and are given in Chapter 5.

- We develop a semi-analytical approach to finding the hitting density of the McKean–Vlasov equation. In PDE terms, this leads to a nonlinear and nonlocal parabolic equation. Using the method of heat potentials, we derive an equivalent coupled system of Volterra integral equations and solve it numerically, or by expansion for a small interaction parameter α. We confirm empirically the convergence of order 1 of the numerical method. The results, submitted for a publication and available in Lipton et al. (2018), are given in Chapter 6.

1.3 Outline of the thesis

In Chapter 2, we present the modeling framework. We start with the main definitions of survival probabilities, credit derivatives, and valuation adjustments. After that, we formulate the general case of the extended default model with mutual liabilities (Lipton (2016)), and show how to formulate corresponding valuation equations for survival probabilities and credit derivatives. As examples, we consider special cases of one, two, and three banks, describe the model and problems for these cases. Finally, under several

additional assumptions we compute the mean-field limit of the model when the number of banks goes to infinity.

In Chapter 3, we consider the case when the asset values follow a simple diffusion process. In this case, for a small number of banks, several quantities are available analytically. We start with a formulation of the problem of transition probability of Brownian motion in $\mathbb{R}_+^N$ and show how it is related to the computation of main characteristics of the model. After that, we review how to compute the transition probability of Brownian motion for $N = 1$ and $N = 2$, which are well-known results in the literature. Next, we develop a novel semi-analytical method for computation of the transition probability for $N = 3$. Finally, we show how to apply these results for the default model with one, two, and three banks.

In Chapter 4, we consider numerical methods for a jump–diffusion model with a small number of banks. We start with the model for one bank. The results for this case are well-known in the literature and are given in Lipton and Sepp (2013). Next, we develop an alternative direction implicit (ADI) type method for the multi-dimensional problem; we prove the stability and deduce the convergence rate, and show how to improve the convergence rate for discontinuous boundary and initial conditions. Finally, we apply the method to the extended default model described in Chapter 2. We show how to compute important model characteristics and calibrate model parameters to the market.

In Chapter 5, we develop a simulation method for the McKean–Vlasov equation arising in a mean-field approximation of the model for independent diffusion processes. We start with a simple particle method, where at each step the loss rate is estimated as the proportion of defaulted banks from the previous step. We show the convergence rate of the method in the number of time steps and the convergence in the number of particles. After that, we improve the convergence rate in the number of time steps by using Brownian bridges. We confirm our results empirically.

In Chapter 6, we solve the McKean–Vlasov equation from the previous chapter by solving the corresponding PDE for the transition probability density. Using the method of heat potentials, we express the solution as a coupled system of Volterra integral equations of the second kind. We propose two methods to solve the system numerically. First, we propose a perturbation method, which uses an expansion in the interaction parameter and works well when it is small. Second, we propose a numerical method based on integral approximation by quadratures. We confirm our results numerically.

In Chapter 7, we offer some conclusions.

Chapter 2

Modeling framework

In this chapter, we give a general overview of the structural default model with mutual obligations developed in Lipton (2016), which is the basis for the analytical and numerical methods developed in the thesis. We start with the main definitions, such as survival probabilities and credit derivatives. Then, we describe the extended default model with mutual liabilities (Lipton (2016)) and derive the main equations.

2.1 Main definitions

In this section, we give the main definitions, which we further apply to the default model. These definitions are model independent. We only assume that we work on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ endowed with filtration $\mathbb{F} = (\mathcal{F})_t)_{0 \leq t \leq T}$, for some $T > 0$, which is large enough to support the main objects of the model (such as firm's asset value process). We also assume that there exists a risk-neutral measure $\mathbb{Q}$ and use it to compute survival probability and price credit derivatives.

2.1.1 Survival probability

In the case of a single bank, its survival probability at time t is the probability of no default between time t and T, and can be written as

$$Q(t,T) = \mathbb{E}^{\mathbb{Q}}[\mathbb{1}_{\{\tau > T\}} \mid \mathcal{F}_t], \tag{2.1}$$

where τ is the default time.

In the case of several banks, the survival probability also depends on the assets of other banks. In this case, we can define joint and marginal survival probabilities as follows.

The joint survival probability of N banks is the probability that none of the banks defaults between t and T. This can be written as

$$Q(t,T) = \mathbb{E}^{\mathbb{Q}}[\mathbb{1}_{\{\tau_1 > T, \ldots, \tau_N > T\}} \mid \mathcal{F}_t], \tag{2.2}$$

where τ_i is the default time of bank i.

On the other hand, the marginal survival probability of bank i is the probability of no default of bank i between t and T conditional on its and all other banks' assets. It can be written as

$$q_i(t,T) = \mathbb{E}^{\mathbb{Q}}[\mathbb{1}_{\{\tau_i > T\}} \mid \mathcal{F}_t], \tag{2.3}$$

where τ_i is the default time of bank i.

2.1.2 Credit Default Swap

Credit default swaps (CDS) are some of the most liquid credit derivatives, which are designed to exchange credit risk of a Reference Name (RN) between a Protection Buyer (PB) and a Protection Seller (PS). PB makes periodic coupon payments to PS conditional on no default of RN up to the nearest payment date, in the exchange for receiving loss given RN's default from PS.

For a CDS contract we define payment dates $t_1 < t_2 < \ldots < t_m$, when PB makes coupon payments c. Most liquid CDSs have maturity 3–5 years and the coupon is usually paid quarterly.

In order to price a CDS contract, usually two legs are considered, fixed $V^{fixed}(t,T)$ and floating $V^{floating}(t,T)$, where the fixed leg accounts for the coupon payments, while the floating leg accounts for the final payment in case of default. In this section, we assume that there is no counterparty risk, i.e. PS and PB cannot default. Mathematically, the values can be written as

$$V^{fixed}(t,T) = \mathbb{E}^{\mathbb{Q}}\left[\sum_{i=1}^{m} D(t,t_i)\mathbb{1}_{\{t_i \le \tau\}}(t_i - t_{i-1}) \mid \mathcal{F}_t\right], \tag{2.4}$$

$$V^{floating}(t,T) = \mathbb{E}^{\mathbb{Q}}\left[(1-R)D(t,\tau)\mathbb{1}_{\{\tau \le T\}} \mid \mathcal{F}_t\right], \tag{2.5}$$

where τ is the default time of the RN, $D(t,s)$ is the discount factor, assumed deterministic recovery rate R, a percentage of defaulted debt which can be recovered.

Taking expectations in the last two equations, we get

$$V^{fixed}(t,T) = \sum_{i=1}^{m} D(t,t_i)Q(t,t_i)(t_i - t_{i-1}), \tag{2.6}$$

$$V^{floating}(t,T) = -(1-R)\int_t^T D(t,s)\,\mathrm{d}Q(t,s), \tag{2.7}$$

where $Q(t,s)$ is the survival probability defined in (2.1).

As a result, the arbitrage-free value of a CDS is given by

$$V^{CDS}(t,T) = V^{floating}(t,T) - cV^{fixed}(t,T). \tag{2.8}$$

Alternatively, a par-spread is also considered, i.e. the CDS spread, when the arbitrage-free value of a CDS contract in (2.8) is zero:

$$c^{par}(t,T) = \frac{V^{floating}(t,T)}{V^{fixed}(t,T)}. \tag{2.9}$$

In the following, for simplicity, we assume that the coupon is paid continuously. Then, the fixed leg can be rewritten as

$$V^{fixed}(t,T) = \int_t^T D(t,s)Q(t,s)\,\mathrm{d}s. \tag{2.10}$$

We also assume that the interest rate is constant. Hence, the discount factor can be written as $D(t,s) = \mathrm{e}^{-r(s-t)}$, where r is the interest rate. In this case, integrating (2.7) by parts, we get

$$\begin{aligned} V^{floating}(t,T) &= (1-R)\left[1 - \mathrm{e}^{-r(T-t)}Q(t,T) - r\int_t^T \mathrm{e}^{-r(s-t)}Q(t,s)\,\mathrm{d}s\right] \\ &= (1-R)\left[1 - \mathrm{e}^{-r(T-t)}Q(t,T) - rV^{fixed}(t,T)\right]. \end{aligned} \tag{2.11}$$

Using the last expression for $V^{floating}(t,T)$, we get the final expression of (2.8) for valuation of a CDS contract without counterparty risk:

$$V^{CDS}(t,T) = -\left(c + r(1-R)\right)V^{fixed}(t,T) + (1-R)\left(1 - \mathrm{e}^{-r(T-t)}Q(t,T)\right). \tag{2.12}$$

Moreover, assuming the interest rate to be zero, (2.12) simplifies to

$$V^{CDS}(t,T) = -cV^{fixed}(t,T) + (1-R)\left(1 - Q(t,T)\right). \tag{2.13}$$

The last formula is quite intuitive. We pay the fixed leg V^{fixed}, and get back the recovery with the default probability $1 - Q(t,T)$.

2.1.3 First-to-Default Swap

A First-to-Default Swap (FTD swap) is a contract designed to exchange credit risk of a basket of Reference Names between a Protection Buyer and a Protection Seller. Similar to a regular CDS, the Protection Buyer (PB) pays a regular coupon payment c to the Protection Seller (PS) up to the first default of any of RNs in the basket or the maturity time T. In return, PS compensates PB for the loss caused by the first default. It can be seen as an extension of a standard CDS to multiple names. Hence, the valuation of an FTD is similar to a regular CDS, where the main difference is that we have to model the joint distribution of the default times of the basket of the RNs, which leads to a higher dimensional problem. Again, here we assume that there is no counterparty risk.

By considering the basket of N reference names, and τ_i, the default time for RN i, for $1 \le i \le N$. Then, similar to (2.4)–(2.5), we can define the fixed and floating legs for an FTD swap

$$V_{FTD}^{fixed}(t,T) = \mathbb{E}^{\mathbb{Q}}\left[\sum_{i=1}^{m} D(t,t_i)\mathbb{1}_{\{t_i \le \tau\}}(t_i - t_{i-1}) \mid \mathcal{F}_t\right], \tag{2.14}$$

$$V_{FTD}^{floating}(t,T) = \mathbb{E}^{\mathbb{Q}}\left[\sum_{i=1}^{N} (1-R_i)D(t,\tau_i)\mathbb{1}_{\{\tau_i \le T, \tau = \tau_i\}} \mid \mathcal{F}_t\right], \tag{2.15}$$

where $\tau = \min_{1\le i\le N}\tau_i$, $D(t,s)$ is the discount factor, and R_i is the recovery rate of RN i.

Taking expectations in (2.14)–(2.15), we get

$$V_{FTD}^{fixed}(t,T) = \sum_{i=1}^{m} D(t,t_i)Q(t,t_i)(t_i - t_{i-1}), \tag{2.16}$$

$$V_{FTD}^{floating}(t,T) = -\sum_{k=1}^{N}(1-R_k)\int_t^T D(t,s)\,\mathrm{d}q_k(t,s), \tag{2.17}$$

where $q_k(t,s)$ is the marginal survival probability of bank k defined in (2.3).

As a result, the arbitrage-free value of an FTD is given by

$$F(t,T) = V_{FTD}^{floating}(t,T) - cV_{FTD}^{fixed}(t,T). \tag{2.18}$$

2.1.4 Credit Default Swap with counterparty risk

In Section 2.1.2 we assumed that there is no counterparty risk, i.e. PS and PB are risk-free. In this section, we show how to valuate a CDS, when counterparties might default. In this section we assume that either one counterparty is risky, or both counterparties are risky.

PS is risky and PB is non-risky Assume that PS is risky, while PB is non-risky. For convenience, we denote all cashflows by

$$CF(t,T) = (1-R_{RN})D(t,\tau)\mathbb{1}_{\{\tau_{RN}\le T\}} - c\sum_{i=1}^{m} D(t,t_i)\mathbb{1}_{\{t_i\le\tau_{RN}\}}(t_i - t_{i-1}). \tag{2.19}$$

Then, the arbitrage-free value of a CDS with counterparty risk is given by

$$\begin{aligned} V_{CPrisk}^{CDS}(t,T) = {} & \mathbb{E}^{\mathbb{Q}}\left[CF(t,T)\mathbb{1}_{\{\tau_{PS}>\min(\tau_{RN},T)\}} \mid \mathcal{F}_t\right] \\ & + \mathbb{E}^{\mathbb{Q}}\left[\left[CF(t,\tau_{PS}) + D(t,\tau_{PS})(R_{PS}V_+^{CDS}(t,\tau_{PS}) + V_-^{CDS}(t,\tau_{PS}))\right]\times\right. \\ & \left.\mathbb{1}_{\{\tau_{PS}\le\min(\tau_{RN},T)\}} \mid \mathcal{F}_t\right], \end{aligned} \tag{2.20}$$

where V^{CDS} is the CDS spread with non-risky counterparties, $V_+ = \max(V,0), V_- = \min(V,0)$, τ_{PS} and τ_{RN} are the default times of PS and RN correspondingly.

PS is non-risky and PB is risky Assume that PB is risky, while PS is non-risky. Similar to the previous section, the arbitrage-free value of a CDS is given by

$$\begin{aligned} V_{CPrisk}^{CDS}(t,T) = {} & \mathbb{E}^{\mathbb{Q}}\left[CF(t,T)\mathbb{1}_{\{\tau_{PB}>\min(\tau_{RN},T)\}} \mid \mathcal{F}_t\right] \\ & + \mathbb{E}^{\mathbb{Q}}\left[\left[CF(t,\tau_{PB}) + D(t,\tau_{PB})(R_{PB}V_-^{CDS}(t,\tau_{PB}) + V_+^{CDS}(t,\tau_{PB}))\right]\times\right. \\ & \left.\mathbb{1}_{\{\tau_{PB}\le\min(\tau_{RN},T)\}} \mid \mathcal{F}_t\right], \end{aligned} \tag{2.21}$$

where $CF(t,T)$ is from (2.19), V^{CDS} is the spread of a CDS with non-risky counterparties, $V_+ = \max(V,0), V_- = \min(V,0)$, τ_{PB} and τ_{RN} are the default times of PB and RN correspondingly.

Both counterparties are risky Now assume both PS and PB are risky. Similar to the previous section denote the cashflows (2.19) by $CF(t,T)$ and $\tau = \min(\tau_{PS}, \tau_{PB})$. Then, the we can compute the value of a CDS

$$\begin{aligned} V^{CDS}_{CPrisk}(t,T) = \mathbb{E}^{\mathbb{Q}}\left[CF(t,T)\mathbb{1}_{\{\tau>\min(\tau_{RN},T)\}} \mid \mathcal{F}_t\right] \\ + \mathbb{E}^{\mathbb{Q}}\left[\left[CF(t,\tau_{PS}) + D(t,\tau_{PS})(R_{PS}V^{CDS}_{+}(t,\tau_{PS}) + V^{CDS}_{-}(t,\tau_{PS}))\right] \times \right. \\ \left. \mathbb{1}_{\{\tau=\tau_{PS}\leq\min(\tau_{RN},T)\}} \mid \mathcal{F}_t\right] \\ + \mathbb{E}^{\mathbb{Q}}\left[\left[CF(t,\tau_{PB}) + D(t,\tau_{PB})(R_{PB}V^{CDS}_{-}(t,\tau_{PB}) + V^{CDS}_{+}(t,\tau_{PB}))\right] \times \right. \\ \left. \mathbb{1}_{\{\tau=\tau_{PB}\leq\min(\tau_{RN},T)\}} \mid \mathcal{F}_t\right]. \quad (2.22) \end{aligned}$$

where V^{CDS} is the spread of a CDS with non-risky counterparties, $V_+ = \max(V,0), V_- = \min(V,0)$, τ_{PS}, τ_{PB}, and τ_{RN} are the default times of PS, PB, and RN correspondingly.

2.1.5 Credit and Debt Value Adjustments

Credit Value Adjustment (CVA) represents the additional costs associated with the possibility of default of the protection seller (PS). Similar, one can consider Debt Value Adjustment, which represents the additional costs associated with the possibility of default of the protection buyer (PB).

There are two possible formulations of CVA and DVA problems. First, we consider unilateral CVA/DVA, where we assume that one of the counterparties is risky and another one is non-risky. Next, we consider bilateral CVA/ DVA, where both counterparties are risky.

Unilateral CVA/DVA We can define unilateral CVA as

$$V^{CVA}(t,T) = V^{CDS}(t,T) - V^{CDS}_{CPrisk}(t,T), \quad (2.23)$$

where $V^{CDS}(t,T)$ is defined in (2.8) and $V^{CDS}_{CPrisk}(t,T)$ is defined in (2.20).

Simple computations yield

$$V^{CVA} = (1-R_{PS})\mathbb{E}^{\mathbb{Q}}[\mathbb{1}_{\{\tau^{PS}<\min(T,\tau^{RN})\}}(V^{CDS}_{t,\tau^{PS}})_+ \,|\mathcal{F}_t], \quad (2.24)$$

Similar, for DVA

$$V^{DVA}(t,T) = V^{CDS}(t,T) - V^{CDS}_{CPrisk}(t,T), \quad (2.25)$$

where $V^{CDS}(t,T)$ is defined in (2.8) and $V^{CDS}_{CPrisk}(t,T)$ is defined in (2.21). It yields,

$$V^{DVA} = (1-R_{PB})\mathbb{E}^{\mathbb{Q}}[\mathbb{1}_{\{\tau^{PB}<\min(T,\tau^{RN})\}}(V^{CDS}_{\tau^{PB}})_- \,|\mathcal{F}_t], \quad (2.26)$$

Bilateral CVA/DVA For bilateral CVA/DVA we assume that both counterparties (PB and PS) are risky. Then

$$V^{CDS}(t,x) = V^{CDS}(t,x) - V^{CDS}_{CPrisk}(t,x), \quad (2.27)$$

where $V^{CDS}(t,x)$ is defined in (2.8) and $V^{CDS}_{CPrisk}(t,x)$ is defined in (2.22).

2.2 Extended default model with mutual liabilities

In the remainder of the chapter we describe the default model from Lipton (2016) and formulate the problems which we further consider in the thesis.

Consider a network of N banks with the i-th bank having directly held assets A_i, external liabilities L_i, and mutual obligations with other banks $L_{ij}, j = 1, \dots N$.

We denote by A_{-i} the interbank assets of bank i and by L_{-i} for interbank liabilities of bank i,

$$A_{-i} = \sum_{j \neq i} L_{ji}, \quad L_{-i} = \sum_{j \neq i} L_{ij}.$$

Then, the total assets and total liabilities can be expressed as

$$\tilde{A}_i = A_i + A_{-i}, \tag{2.28}$$

$$\tilde{L}_i = L_i + L_{-i}. \tag{2.29}$$

Accordingly, an individual bank's capital is given by

$$E_i = \tilde{A}_i - \tilde{L}_i = A_i + A_{-i} - L_i - L_{-i}. \tag{2.30}$$

We use the matrix notation

$$\begin{aligned} \boldsymbol{A} &= (A_{ij}), \quad A_{ii} = A_i, \quad A_{ij} = L_{ji}, \\ \boldsymbol{L} &= (L_{ij}), \quad L_{ii} = L_i, \quad L_{ij} = L_{ij}. \end{aligned} \tag{2.31}$$

2.2.1 Dynamics of assets and liabilities

We assume that assets are driven by jump-diffusion processes with negative exponential jumps as described in Lipton and Sepp (2013),

$$\frac{\mathrm{d}A_i}{A_i} = (r - \kappa_i \lambda_i(t))\mathrm{d}t + \sigma_i \mathrm{d}W_i(t) + (\mathrm{e}^{J_i} - 1)\mathrm{d}N_i(t), \tag{2.32}$$

where r is growth rate, σ_i are corresponding volatilities, W_i are correlated Brownian motions,

$$\mathrm{d}W_i(t)\mathrm{d}W_j(t) = \rho_{ij}\mathrm{d}t, \tag{2.33}$$

N_i are Poisson processes independent of W_i, λ_i are intensities of jump arrivals, J_i are random jump sizes with probability density function

$$\tilde{\omega}_i(s) = \begin{cases} 0, & s > 0, \\ \vartheta_i \mathrm{e}^{\vartheta_i s}, & s \leq 0, \end{cases} \tag{2.34}$$

with parameters $\vartheta_i > 0$; and κ_i are jump compensators

$$\kappa_i = \mathbb{E}^{\mathbb{Q}}[\mathrm{e}^{J_i} - 1] = -\frac{1}{\vartheta_i + 1}. \tag{2.35}$$

The jump processes are correlated in the spirit of Marshall and Olkin (1967). Denote by $\Pi^{(N)}$ the set of all subsets of N names except the empty subset. Then, for every subset $\pi \in \Pi^{(N)}$ we associate the Poisson process $N_\pi(t)$ with intensity λ_π:

- $N_\pi(0) = 0$,

- a process with independent increments,

- $\mathbb{P}(N_\pi(t) = k) = \frac{(\lambda_\pi t)^k}{k!} \mathrm{e}^{-\lambda_\pi t}$.

We define the process $N_i(t)$ as

$$N_i(t) = \sum_{\pi \in \Pi^{(N)}} \mathbb{1}_{\{i \in \pi\}} N_\pi(t) \tag{2.36}$$

with

$$\lambda_i = \sum_{\pi \in \Pi^{(N)}} \mathbb{1}_{\{i \in \pi\}} \lambda_\pi. \tag{2.37}$$

Using this jump definition, we assume that there are both systemic and idiosyncratic sources of jumps. In practice, it is sufficient to consider $N + 1$ sets of $\Pi^{(N)}$: $\{i\}, i = 1, \ldots, N$ and $\{1, \ldots, N\}$. To exclude other subsets, we can put their intensities $\lambda_\pi = 0$. If we need more risk factors, we can include additional subsets, for example, representing particular regions or sectors.

A special case of asset's dynamics without jumps can also be considered. In this case, we set all intensities λ_π to 0,

$$\frac{\mathrm{d}A_i}{A_i} = r\mathrm{d}t + \sigma_i \mathrm{d}W_i(t). \tag{2.38}$$

We assume that liabilities are deterministic and grow with rate r,

$$\frac{\mathrm{d}L_i}{L_i} = r\,\mathrm{d}t, \quad \frac{\mathrm{d}L_{ij}}{L_{ij}} = r\,\mathrm{d}t. \tag{2.39}$$

In the following, we assume r to be 0, in this case, liabilities are constant.

2.2.2 Default boundaries

Bank i defaults if its asset value process crosses its default boundary. The default time for bank i is

$$\tau_i = \inf\{t \mid A_i(t) \leq \Lambda_i\}, \quad i = 1, \ldots, N, \tag{2.40}$$

and we also define $\tau = \min_{1 \leq i \leq N} \tau_i$.

When $t < \tau$, the default boundary Λ_i for bank i is

$$A_i \leq \Lambda_i := \begin{cases} R_i \tilde{L}_i - A_{-i} \equiv \Lambda_i^<, & t < T, \\ \tilde{L}_i - A_{-i} \equiv \Lambda_i^=, & t = T, \end{cases} \tag{2.41}$$

where $0 \leq R_i \leq 1$ is the recovery rate.

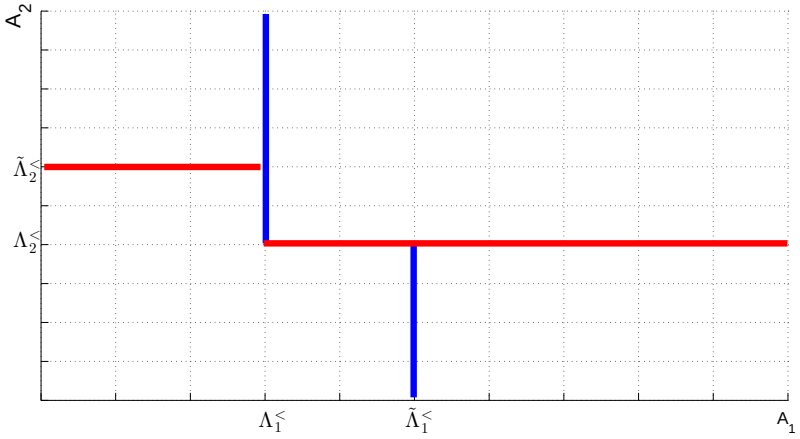


Figure 2.1: Example of default boundaries for two banks at $t < T$.

If the k-th bank defaults at intermediate time t' (i.e. $\tau_k < t$), then we consider a system of $N-1$ banks and their modified default boundaries. We re-numerate surviving banks by applying the function

$$i \to i' = \phi^k(i) = \begin{cases} i, & i < k, \\ i-1, & i > k. \end{cases} \tag{2.42}$$

We also introduce the inverse function

$$i \to i' = \psi^k(i) = \begin{cases} i, & i < k, \\ i+1, & i \geq k. \end{cases} \tag{2.43}$$

The corresponding asset and liability matrices have the form

$$\begin{aligned} &\boldsymbol{A}^{(k)}(t) = \left(A_{ij}^{(k)}(t)\right), \quad \boldsymbol{L}^{(k)}(t) = \left(L_{ij}(t)\right), \\ &A_{ii}^{(k)}(t) = A_{\psi^k(i)\psi^k(i)}(t), \quad A_{ij}^{(k)}(t) = A_{\psi^k(i)\psi^k(j)}(t), \\ &L_{ii}^{(k)}(t) = L_{\psi^k(i)\psi^k(i)}(t) - L_{\psi^k(i)k}(t') + R_k L_{k\psi^k(i)}(t'), \quad L_{ij}^{(k)}(t) = L_{\psi^k(i)\psi^k(j)}(t), \\ &t > t', i, j = 1, \ldots, N-1. \end{aligned} \tag{2.44}$$

The corresponding default boundaries are modified to

$$A_i \leq \Lambda_i^{(k)} = \begin{cases} R_{\psi^{(k)}(i)} \tilde{L}_i^{(k)} - A_{-i}^{(k)}, & t < T, \\ \tilde{L}_i^{(k)} - A_{-i}^{(k)}, & t = T. \end{cases} \tag{2.45}$$

It is clear that

$$\Delta\Lambda_i^{(k)} = \Lambda_i^{(k)} - \Lambda_i = \begin{cases} (1 - R_i R_k) A_{-i}^{(k)}, & t < T, \\ (1 - R_k) A_{-i}^{(k)}, & t = T. \end{cases} \tag{2.46}$$

Thus, $\Delta\Lambda_i^{(k)} > 0$ and the corresponding default boundaries move to the right.

We illustrate this effect on the example of two banks in Figure 2.1.

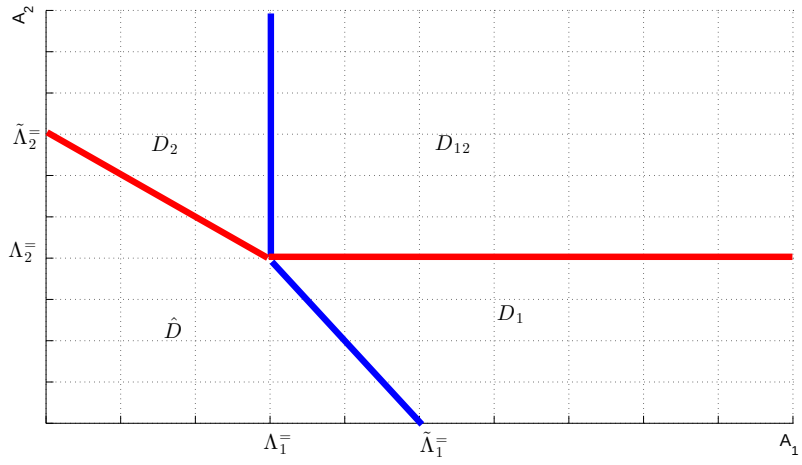


Figure 2.2: Example of default boundaries for two banks at $t = T$.

2.2.3 Terminal conditions

We need to describe the settlement process at time $t = T$. We shall do it in the spirit of Eisenberg and Noe (2001). Since at time T full settlement is expected, we assume that bank i will pay the fraction ω_i of its total liabilities to creditors. It means that if $\omega_i = 1$, the bank pays all liabilities (both external and interbanking) and survives. On the other hand, if $0 < \omega_i < 1$, it means that the bank i defaults, and pays only a fraction of liabilities. Thus, we can describe the terminal condition as a system of equations

$$\min\left\{A_i(T) + \sum_{j\neq i}\omega_j L_{ji}, L_i + \sum_{j\neq i} L_{ij}\right\} = \omega_i\left(L_i + \sum_{j\neq i} L_{ij}\right). \tag{2.47}$$

Dividing both parts by $L_i + \sum_{j\neq i} L_{ij}$, we have

$$\min\left\{\frac{A_i(T) + \sum_{j\neq i}\omega_j L_{ji}}{L_i + \sum_{j\neq i} L_{ij}}, 1\right\} = \omega_i. \tag{2.48}$$

Denote the left-hand side as $\Phi(\omega)$. Then, (2.48) can be rewritten as

$$\Phi(\omega) = \omega. \tag{2.49}$$

Eisenberg and Noe (2001) have shown that the equation (2.49) is a non-expanding mapping in the standard Euclidian metric and formulated conditions under which there is just one fixed point. We assume that these conditions are satisfied and for each $A(T)$ there is a unique vector $\omega = (\omega_1, \ldots, \omega_N)^T$ such that the condition (2.48) is satisfied. In Figure 2.2, we show default boundaries on the example of two banks.

The actual terminal condition depends on the problem we want to solve. Based on the vector ω we can find which banks defaulted, $\omega_i < 1$, and which did not, $\omega_i = 1$, and, therefore, determine the terminal condition. In the following, we consider several useful examples of the terminal conditions.

Denote the vector $I = (I_1, \ldots, I_N)^T$ with elements $I_i \in \{0, 1\}$. Denote by $\mathcal{D}(I)$ the following domain:

$$\mathcal{D}(I) = \{A(T) \mid \lfloor \omega_i(A(T)) \rfloor = I_i\}. \tag{2.50}$$

In this domain, the bank i survives if $I_i = 1$ and defaults otherwise. For example, in the domain $\mathcal{D}(1, \ldots, 1)$ all banks survive, and in the domain $\mathcal{D}(0, 1, \ldots, 1)$ the first bank defaults and all other banks survive.

Now we can formulate the terminal conditions for joint and marginal survival probabilities. For the joint survival probability (2.2), the terminal condition has the form

$$Q(T, T, A) = \mathbb{1}_{\{A \in \mathcal{D}(1,\ldots,1)\}}. \tag{2.51}$$

For the marginal survival probability of bank i (2.3), the terminal condition is

$$q_i(T, T, A) = \mathbb{1}_{\left\{A \in \cup_{I^{(i)}} \mathcal{D}(I^{(i)})\right\}}, \tag{2.52}$$

where $I^{(i)}$ is the set of indicator vectors where the i^{th} position equals to one.

We shall discuss the terminal conditions of more complicated products later in this chapter.

2.2.4 Formulation of backward Kolmogorov equation

Introduce normalized dimensionless variables

$$\bar{t} = \Sigma^2 t, \quad X_i = \frac{\Sigma}{\sigma_i} \ln\left(\frac{A_i}{\Lambda_i^<}\right), \quad \bar{\lambda}_i = \frac{\lambda_i}{\Sigma^2}, \quad \bar{T} = \Sigma^2 T, \tag{2.53}$$

where

$$\Sigma = \left(\prod_{i=1}^{N} \sigma_i\right)^{1/N}.$$

Denote also

$$\xi_i = -\left(\frac{\sigma_i}{2\Sigma} + \kappa_i \bar{\lambda}_i\right), \quad \zeta_i = \frac{\Sigma}{\sigma_i}. \tag{2.54}$$

$X_i(t)$ is also called the distance-to-default of bank i. Applying Itô's formula for X_i we find its dynamics

$$\mathrm{d}X_i = \xi_i \mathrm{d}\bar{t} + \mathrm{d}W_i(\bar{t}) + \zeta_i J_i \mathrm{d}N_i(\bar{t}). \tag{2.55}$$

The default boundaries change to

$$X_i \le \mu_i := \begin{cases} \mu_i^< = 0, & t < T, \\ \mu_i^= = \frac{\Sigma}{\sigma_i} \ln\left(\frac{\Lambda_i^=(\bar{t})}{\Lambda_i^<(\bar{t})}\right), & t = T. \end{cases} \tag{2.56}$$

We further omit bars for brevity and for convenience denote $X(t) = (X_1(t), \ldots, X_N(t))^T$.

Consider the terminal condition $\psi(X)$, intermediate contract payment (for example, coupon payment) $\chi(s, x)$, the boundary conditions $\phi_{i,0}$ (the payoff in case of intermediate

default of bank i). Then, the arbitrage-free value at time t is given by

$$V(x,t) = \mathbb{E}^{\mathbb{Q}}\left[\psi(X(T))\mathbb{1}_{\{\tau\geq T\}} + \int_t^T \chi(s,X(s))\mathbb{1}_{\{\tau>s\}}\,\mathrm{d}s \right. \\ \left. \sum_{i=1}^{N} \phi_{i,0}(\tau_i, X_{-i}(\tau_i))\mathbb{1}_{\{\tau_i<T,\tau_i=\tau\}} \mid X(t)=x\right], \quad (2.57)$$

where $X_{-i} = (X_1,\ldots,X_{i-1},X_{i+1},\ldots,X_N)^T$.

According to the Feynman–Kac formula, the corresponding pricing equation is the Kolmogorov backward equation

$$\frac{\partial V}{\partial t} + \mathcal{L}^{(N)}V = \chi(t,x), \quad (2.58)$$

$$V(t,x_{0,k}) = \phi_{0,k}(t,x_{-k}), \quad V(t,x) \underset{x_k\to+\infty}{\longrightarrow} \phi_{\infty,k}(t,x_{-k}), \quad (2.59)$$

$$V(T,x) = \psi(x), \quad (2.60)$$

where

$$\begin{aligned} x &= (x_1,\ldots,x_N),\\ x_{0,k} &= (x_1,\ldots,\underset{k}{0},\ldots,x_N),\\ x_{-k} &= (x_1,\ldots,x_{k-1},x_{k+1},\ldots,x_N), \end{aligned} \quad (2.61)$$

with Kolmogorov backward operator

$$\mathcal{L}^{(N)}f = \frac{1}{2}\sum_{i,j=1}^{N}\rho_{ij}f_{x_ix_j} + \sum_{i=1}^{N}\xi_i f_{x_i} + \sum_{\pi\in\Pi^{(N)}}\lambda_\pi\prod_{i\in\pi}\mathcal{J}_i f(x) - \sum_{\pi\in\Pi^{(N)}}\lambda_\pi f(x) = \\ = \frac{1}{2}\Delta_\rho f + \xi\cdot\nabla f + \mathcal{J}f - vf, \quad (2.62)$$

where

$$\mathcal{J}f(x) = \sum_{\pi\in\Pi^{(N)}}\lambda_\pi\prod_{i\in\pi}\mathcal{J}_i f(x), \quad (2.63)$$

$$\mathcal{J}_i f(x) = \varsigma_i\int_0^{x_i} f(x_1,\ldots,x_i-u,\ldots,x_N)\mathrm{e}^{-\varsigma_i u}\,\mathrm{d}u, \quad (2.64)$$

and $\varsigma_i = \sigma_i\vartheta_i/\Sigma$.

In the remainder of the section we formulate the Kolmogorov backward equation for specific products described in Section 2.1. For all products, the terminal time is T and we omit it for brevity. For example, in Section 2.1 we used $Q(t,T)$ for the survival probability; now we omit T, but include the dependence on the initial state x. Hence, the survival probability is denoted by $Q(t,x)$.

2.2.5 Joint survival probability

We defined the joint survival probability in Section 2.1.1. Adapting it to the framework described in this section, we get

$$Q(t,x) = \mathbb{E}^{\mathbb{Q}}[\mathbb{1}_{\{X(T)\in\mathcal{D}(1,\ldots,1)\}} \mid X(t)=x]. \quad (2.65)$$

Then, applying (2.58)–(2.60) with $\psi(x) = \mathbb{1}_{\{X(T)\in\mathcal{D}(1,\dots,1)\}}, \chi(t,x) = 0, \phi_{i,0}(t,y) = 0$, we get

$$\begin{aligned} &\frac{\partial Q}{\partial t} + \mathcal{L}^{(N)}Q = 0, \\ &Q(t, x_{0,k}) = 0, \\ &Q(T,x) = \mathbb{1}_{\{x\in\mathcal{D}(1,\dots,1)\}}. \end{aligned} \tag{2.66}$$

2.2.6 Marginal survival probability

The marginal survival probability for bank i was defined in Section 2.1.1. In our framework, it becomes

$$q_i(t,x) = \mathbb{E}^{\mathbb{Q}}\left[\mathbb{1}_{\left\{X(T)\in\cup_{I^{(i)}}\mathcal{D}(I^{(i)}),\tau\geq T\right\}} + \sum_{j=1}^{N}\phi_{j,0}(\tau_j, X_{-j}(\tau_j))\mathbb{1}_{\{\tau_j<T,\tau_i=\tau\}} \mid X(t) = x\right], \tag{2.67}$$

where $I^{(i)}$ is the set of indicator vectors, such that $I_i = 1$,

$$X_{-i}(t) = (X_1(t), \dots, X_{i-1}(t), X_{i+1}(t), \dots, X_N(t))^T, \tag{2.68}$$

and $\phi_{j,0}(t,y)$ is the survival probability for the system of $N-1$ banks, when bank j defaulted. Correspondingly, $\phi_{i,0}(t,y) = 0$.

Applying (2.58)–(2.60) with $\psi(x) = \mathbb{1}_{\left\{X(T)\in\cup_{I^{(i)}}\mathcal{D}(I^{(i)})\right\}}, \chi(t,x) = 0$, we get

$$\begin{aligned} &\frac{\partial}{\partial t}q_i(t,x) + \mathcal{L}^{(N)}q_i(t,x) = 0, \\ &q_i(t,x_{0,i}) = 0, \quad q_i(t,x_{0,k}) = \Xi(t,x_{-k}) = \begin{cases}\chi_{ik,0}(t,x_{-k}), & x_i \geq \tilde{\mu}_i, \\ 0, & x_i < \tilde{\mu}_i.\end{cases}, \\ &q_i(t,x) \underset{x_k\to+\infty}{\longrightarrow} 1, \quad q_i(t,x) \underset{x_k\to+\infty}{\longrightarrow} \chi_{ik,\infty}(t,x_{-k}), \\ &q_i(T,x) = \mathbb{1}_{\left\{x\in\cup_{I^{(i)}}\mathcal{D}(I^{(i)})\right\}}. \end{aligned} \tag{2.69}$$

where $k \neq i$, $\chi_{ik,0}(t,x_{-k})$ and $\chi_{ik,\infty}(t,x_{-k})$ are the solutions of $N-1$ dimensional problems with corresponding default boundaries (for $\chi_{i,0}(t,x_{-k})$ the i-th bank defaults, for $\chi_{i,\infty}(t,x_{-k})$ it does not).

2.2.7 Credit default swap

Consider a CDS contract written on the first bank without counterparty risk. Denote its spread $C_1(t,x)$.[1] As in Section 2.1.2, we assume that the coupon is paid continuously and equals c. Then, in our framework, (2.13) becomes

$$\begin{aligned} C_1(t,x) = \mathbb{E}^{\mathbb{Q}}\Bigg[&-c\int_t^T \mathbb{1}_{\{\tau>s\}}\,\mathrm{d}s + (1-\min(R_1,\tilde{R}_1)\mathbb{1}_{\{X(T)\in\cup_{i_2,\dots,i_N\in(0,1)}\mathcal{D}(0,i_2,\dots,i_N)\}} \\ &+(1-R_1)\mathbb{1}_{\{\tau_1<T,\tau_1=\tau\}} + \sum_{i=2}^{N} c_{1i,0}(\tau_i, X_{-i}(\tau_i))\mathbb{1}_{\{\tau_i<T,\tau_i=\tau\}} \mid X(t) = x\Bigg], \end{aligned} \tag{2.70}$$

[1]For the CDS contracts written on the bank i, the similar expression could be provided by analogy.

where

$$X_{-i}(t) = (X_1(t), \dots, X_{i-1}(t), X_{i+1}(t), \dots, X_N(t))^T, \tag{2.71}$$

$$\tilde{R}_1 = \min\left[1, \frac{A_1(T) + \sum_{i=2}^{N} \omega_i L_{i1}(T)}{L_1(T) + \sum_{i=2}^{N} \omega_i L_{1i}(T)}\right]. \tag{2.72}$$

Here, $\omega = \omega(X_T)$ is determined from (2.48), and $c_{1i,0}(t,y)$ is the CDS spread for the system of $N-1$ banks, when bank i defaulted.

We thus apply (2.58)–(2.60) with $\chi(t,x) = c$ and terminal condition

$$\psi(x) = (1 - \min(R_1, \tilde{R}_1)) \cdot \mathbb{1}_{x \in \cup_{i_2,\dots,i_N \in (0,1)} \mathcal{D}(0,i_2,\dots,i_N)},$$

where $\omega = \omega(x)$ is defined in (2.48).

Thus, the pricing problem for CDS contract on the first bank is

$$\begin{aligned}
&\frac{\partial}{\partial t} C_1(t,x) + \mathcal{L}^{(N)} C_1(t,x) = c,\\
&C_1(t, x_{0,1}) = 1 - R_1, \quad C_1(t,x) \underset{x_k \to +\infty}{\longrightarrow} c_{1k,\infty}(t, x_{-k}),\\
&C_1(t, x_{0,k}) = \Xi(t, x_{-k}) = \begin{cases} c_{1k,0}(t, x_{-k}), & x_1 \geq \tilde{\mu}_1,\\ 1 - R_1, & x_1 < \tilde{\mu}_1. \end{cases}\\
&C_1(T,x) = (1 - \min(R_1, \tilde{R}_1(\omega))) \cdot \mathbb{1}_{\{x \in \cup_{i_2,\dots,i_N \in (0,1)} \mathcal{D}(0,i_2,\dots,i_N)\}},
\end{aligned} \tag{2.73}$$

where $k \neq 1$, $c_{1k,0}(t, x_{-k})$ and $c_{1k,\infty}(t, x_{-k})$ are the solutions of the corresponding $N-1$ dimensional problems.

2.2.8 First-to-Default Swap

Consider the FTD contract referenced on N banks and denote its price $F(t,x)$. An FTD swap was defined in Section 2.1.3. Assuming the coupon is paid continuously, in our framework (2.18) becomes

$$\begin{aligned}
F(t,x) = \mathbb{E}^{\mathbb{Q}} \Bigg[-c \int_t^T \mathbb{1}_{\{\tau > s\}} \, \mathrm{d}s + \sum_{I=(i_1,\dots,i_N), i_k \in (0,1)} \beta_I \mathbb{1}_{\{X(T) \in \mathcal{D}(I)\}} \\
+ \sum_{i=1}^{N} (1 - R_i) \mathbb{1}_{\{\tau_i < T, \tau_i = \tau\}} \mid X(t) = x \Bigg],
\end{aligned} \tag{2.74}$$

where

$$\beta_I = 1 - \min_{k: I_k = 0} \left[\min\{\tilde{R}_k, R_k\}\right],$$

and

$$\tilde{R}_k = \min\left[1, \frac{A_k(T) + \sum_{i \neq k} \omega_i L_{ik}(T)}{L_k(T) + \sum_{i \neq k} \omega_i L_{ki}(T)}\right]$$

with $\omega = \omega(X_T)$ defined in (2.48).

Applying (2.58)–(2.60) with $\chi(t,x) = c$ and terminal condition

$$\psi(x) = \sum_{I=(i_1,\dots,i_N), i_k \in (0,1)} \beta_I \mathbb{1}_{\{x \in \mathcal{D}(I)\}},$$

we get the pricing problem for an FTD

$$\begin{aligned} &\frac{\partial}{\partial t} F(t,x) + \mathcal{L}^{(N)} F(t,x) = c, \\ &F(t, x_{0,k}) = 1 - R_k, \quad F(t,x) \underset{x_k \to +\infty}{\longrightarrow} f_{k,\infty}(t, x_{-k}), \\ &F(T,x) = \sum_{I=(i_1,\dots,i_N), i_k \in (0,1)} \beta_I \mathbb{1}_{\{x \in \mathcal{D}(I)\}}, \end{aligned} \tag{2.75}$$

where $f_{k,\infty}(t, x_{-k})$ is the solution of the corresponding $N-1$ dimensional problem without the bank k.

2.2.9 Credit and Debt Value Adjustments for a CDS

We associate $X_1(t)$ with a protection seller (PS), $X_2(t)$ with a protection buyer (PB), and $X_3(t)$ with a reference name (RN); all other banks are numbered 4 to N. As we defined in Section 2.1.5, the CVA can be computed as

$$\begin{aligned} V^{CVA}(t,x) = \mathbb{E}^{\mathbb{Q}}[(1-R_1)\mathbb{1}_{\{\tau_1 < T, \tau_1 = \tau\}} (V^{CDS}_{\tau_1})_+ \\ + \sum_{i=4}^{N} \mathbb{1}_{\{\tau_i < T, \tau_i = \tau\}} \tilde{V}^{CVA,i}(\tau_i) \mid X(t) = x], \end{aligned} \tag{2.76}$$

where $\tilde{V}^{CVA,i}$ is CVA for the system of $N-1$ banks with modified boundaries assuming bank i defaulted.

Then, applying (2.58)–(2.60) with $\psi(x) = 0, \chi(t,x) = 0$, we get

$$\begin{aligned} &\frac{\partial}{\partial t} V^{CVA}(t,x) + \mathcal{L}^{(N)} V^{CVA} = 0, \\ &V^{CVA}(t, x_{0,1}) = (1-R_1) V^{CDS}(t, y_1)_+, \\ &V^{CVA}(t, x_{0,2}) = 0, \quad V^{CVA}(t, x_{0,3}) = 0, \\ &V^{CVA}(t, x_{0,k}) = \tilde{V}^{CVA,i}(t, x_{-k}), \quad k \geq 3. \end{aligned} \tag{2.77}$$

Similarly, for the Debt Value Adjustment

$$\begin{aligned} V^{DVA}(t,x) = \mathbb{E}^{\mathbb{Q}}[(1-R_2)\mathbb{1}_{\{\tau_2 < T, \tau_2 = \tau\}} (V^{CDS}_{\tau_2})_- \\ + \sum_{i=4}^{N} \mathbb{1}_{\{\tau_i < T, \tau_i = \tau\}} \tilde{V}^{DVA,i}(\tau_i) \mid X(t) = x], \end{aligned} \tag{2.78}$$

where $\tilde{V}^{DVA,i}$ is DVA for the system of $N-1$ banks with modified boundaries assuming bank i defaulted.

Applying (2.58)–(2.60) with $\psi(x) = 0, \chi(t,x) = 0$, we get

$$\begin{aligned}
&\frac{\partial}{\partial t}V^{DVA}(t,x) + \mathcal{L}^{(N)}V^{DVA} = 0,\\
&V^{DVA}(t,x_{0,2}) = (1-R_2)V^{CDS}(t,y_2)_-,\\
&V^{DVA}(t,x_{0,1}) = 0, \quad V^{DVA}(t,x_{0,3}) = 0,\\
&V^{DVA}(t,x_{0,k}) = \tilde{V}^{DVA,i}(t,x_{-k}), \quad k \geq 3.
\end{aligned} \tag{2.79}$$

In Appendix A we explicitly consider special cases of $N = 1, 2, 3$ of the model.

2.3 Infinite limit of the banking system

As we saw before in (2.58)–(2.60), to solve the model with N banks, we have to solve an N-dimensional PDE/PIDE, which is computationally infeasible for large values of N. Therefore, in this section we consider an infinite-dimensional limit of the banking system under further simplifying assumptions. Specifically, we assume that there is an infinite number of banks, whose values are drawn from the same distribution independently, and we want to model a "typical representative". Thus, instead of the N-dimensional system, we consider a one-dimensional process which takes into account the interaction effect through dependence of the coefficients on its distribution (a McKean-Vlasov process).

In this section, we assume that there are no jumps, and the asset value process of bank i follows (2.38)

$$\frac{\mathrm{d}A_i}{A_i} = r\,\mathrm{d}t + \sigma_i\,\mathrm{d}W_i.$$

We also assume that W_i are independent standard Brownian motions for $1 \leq i \leq N$. It means that interconnection of the system comes only from interbank liabilities.

At time $t = 0$, following Section 2.2.2,

$$\Lambda_i(t) = R_i\left(L_i + \sum_{j\neq i} L_{ij}\right) - \sum_{j\neq i} L_{ji},$$

where $0 < R_i < 1$ is the recovery rate of bank i, i.e., bank i defaults if the recovery value of its liabilities, external and to other banks, exceeds the sum of its assets, external and from other banks.

Since liabilities and recovery rates are assumed constant in time, the default boundary remains constant until some bank defaults. If bank k defaults at time t, the default boundary of bank i jumps by (2.46),

$$\Delta\Lambda_i(t) = (1 - R_iR_k)L_{ki}.$$

In the following, we assume that the banks have the same parameters, i.e., $\sigma_i = \sigma$ and $R_i = R$, for some r, σ, R, $L_i = L$ and $L_{ij} = \frac{\gamma}{N}$, both constant, respectively, for some $L, \gamma > 0$. In particular, this implies that the asset value processes are exchangeable, and we have $\Lambda_i(0) = \Lambda^0$ for some Λ^0, which will allow us to take a large-pool limit.

As a result, we can write $\Lambda_i(t)$ as

$$\Lambda_i(t) = \Lambda_i(0) + \frac{\gamma}{N}\sum_{k\neq i}(1 - R^2)\mathbb{1}_{\{\tau_k \leq t\}}.$$

Considering the change of variables

$$\bar{t} = \sigma^2 t, \quad , X_i(t) = \log(A_i(t)/\Lambda_i(t)),$$

and omitting bars for brevity, we have

$$X_i(t) = X_i(0) + \frac{1}{\sigma^2}(r - \sigma^2/2)t + W_t^i - \log\left(1 + \frac{\gamma}{N}\sum_{k\neq i}(1-R^2)\frac{1}{\Lambda_i(0)}\mathbb{1}_{\{\tau_k \leq t\}}\right).$$

Using the approximation $\log(1+x) \approx x$ for small x (i.e., assuming only a small proportion of the banks have yet defaulted), we get for $t < \tau_i$

$$\begin{aligned} X_i(t) &= X_i(0) + \frac{1}{\sigma^2}(r - \sigma^2/2)t + W_i(t) - \frac{\gamma(1-R^2)}{\Lambda_i(0)}L^N(t), \\ L^N(t) &= \frac{1}{N}\sum_k \mathbb{1}_{\{\tau_k \leq t\}}. \end{aligned}$$

For simplicity, we take the special case $r - \sigma^2/2 = 0$ (in any case, this term will be small for realistic parameter values and not have any qualitative impact on the results).

Then, using propagation of chaos as in Hambly et al. (2018), one can obtain that in the limit for $N \to \infty$ all $X_i(t)$ have the same distribution as X given by[2]

$$\begin{aligned} X(t) &= X_0 + W(t) - \alpha L(t), \\ L(t) &= \mathbb{P}(\tau \leq t), \\ \tau &= \inf\{t \in [0,T] : X(t) \leq 0\}, \end{aligned} \tag{2.80}$$

where $\alpha = \frac{\gamma(1-R^2)}{\Lambda_0}$ and X_0 is a random variable with density f_0.

We consider a right-continuous solution $L(t)$. At the points of discontinuity, we say that the banking system exhibits a "blow-up", which means that there is a chain of defaults, and, as result, the system loses a significant mass (a finite amount at a single time).

The distribution of the stopped process $X_{t\wedge\tau}$ can be written as $L_t\delta + p$, where δ is the atomic measure concentrated at 0 and p is the continuous component. Writing the increasing process L as $\alpha L(t) = -\int_0^t \mu(t')\,dt'$ for some negative μ, p satisfies

$$\begin{aligned} p_t(t,x) &= -\mu(t)\,p_x(t,x) + \frac{1}{2}p_{xx}(t,x), \\ p(0,x) &= f_0(x), \\ p(t,0) &= 0. \end{aligned} \tag{2.81}$$

Using the relation $L(t) = 1 - \int_0^\infty p(t,x;z)\,dx$, we can express μ in terms of p by

$$g(t) := \frac{\mathrm{d}L_t}{\mathrm{d}t} = -\int_0^\infty p_t\,dx = \mu(t)\int_0^\infty p_x\,dx - \frac{1}{2}\int_0^\infty p_{xx}\,dx = \frac{1}{2}p_x(t,0), \tag{2.82}$$

[2]Note the slight ambiguity between the liabilities above and the loss function below, which are both denoted by L. It should be clear from the context and indices applied which one is being referred to, hence for consistency with the literature we keep the notation.

where we have used the PDE (2.81) as well as $p(t,0) = 0$ and $\lim_{x\to\infty} p(t,x) = \lim_{x\to\infty} p_x(t,x) = 0$. Hence, (2.81) can be written in the self-consistent form

$$\begin{aligned} p_t\,(t,x) &= \frac{\alpha}{2} p_x(t,0)\, p_x\,(t,x) + \frac{1}{2} p_{xx}\,(t,x)\,, \\ p\,(0,x) &= f_0(x), \\ p\,(t,0) &= 0. \end{aligned} \tag{2.83}$$

Chapter 3

Analytical methods for the model without jumps

3.1 Introduction

In this chapter, we consider analytical and semi-analytical methods developed to compute various important characteristics of the model. In general, the analytical solution is not available, therefore, we concentrate on several particular cases. We consider the case when there are no jumps, i.e. all jump intensities $\lambda_i = 0$. In this case, the infinitesimal generator $\mathcal{L}^{(N)}$ (2.62) simplifies to

$$\mathcal{L}^{(N)} f = \frac{1}{2} \sum_{i,j=1}^{N} \rho_{ij} f_{x_i x_j} + \sum_{i=1}^{N} \xi_i f_{x_i} = \frac{1}{2} \Delta_\rho f + \xi \cdot \nabla f. \tag{3.1}$$

We solve the pricing problem (2.58)–(2.60) using the Green's function approach. We define the Green's function and give its interpretation as a transition probability of a Markov process in Section 3.2; in Section 3.3, we show how to eliminate the drift term and cross–derivatives. In Section 3.4, we show how to find the solution for the one-dimensional problem, which is a classical problem and solved analytically using the method of images (Lipton (2001)). In Section 3.5, we solve the two-dimensional problem using separation of variables and show that the Green's function for this PDE can be represented as a series of Bessel functions (He et al. (1998), Lipton and Sepp (2009), Itkin and Lipton (2017), Zhou (2001b)). In Section 3.6, we develop a novel semi-analytical method to solve the three-dimensional problem. Using separation of variables, we represent the solution as a double series of special functions, where unknown parameters are found as a solution of a nonlinear eigenvalue problem.

Further, we show the financial applications of the developed methods. In Section 3.7, we show how to compute model characteristics for the model with one bank. This case is classical and fully considered in Lipton and Sepp (2013). In Section 3.8, we consider the model with three banks, and show how to compute survival probabilities and bilateral CVA/DVA.

3.2 Transition probability of Brownian motion in $\mathbb{R}_+^N$

The transition probability of multi-dimensional Brownian motion in $\mathbb{R}_+^N, N \geq 1$, with absorption at the boundary, occurs in different areas of applied mathematics, and mathematical finance in particular. For example, it appears in counterparty credit risk modeling (Lipton and Savescu (2014), Lipton and Sepp (2009), Itkin and Lipton (2017), Zhou (2001b)), market microstructure (Cont and De Larrard (2012), Lipton et al. (2014)), and exotic option pricing (Escobar et al. (2014)) for barrier options on three assets with special correlation structure; He et al. (1998) for lookback options whose payoff depends on the minimum, maximum, and endpoint of the asset price process; Lipton (2001)).

We consider a N-dimensional correlated Brownian motion with constant drift $(\mu_i)_{1\leq i\leq N}$,

$$\mathrm{d}X_i(s) = \mu_i \,\mathrm{d}s + \mathrm{d}W_i(s), \qquad X_i(t) = x_i, \qquad 1 \leq i \leq N,$$

with infinitesimal generator

$$\mathcal{L}^{(N)} = \frac{1}{2}\sum_{i,j=1}^{N} \rho_{ij} \frac{\partial^2}{\partial x_i \partial x_j} + \sum_{i=1}^{N} \mu_i \frac{\partial}{\partial x_i},$$

where $P = (\rho_{ij})_{1\leq i,j\leq N}$ is a symmetric positive definite matrix with $\rho_{ii} = 1$, $1 \leq i \leq N$, the correlation matrix of the standard Brownian motion W.

Now consider the stopped process $X(s \wedge \tau)$, with $\tau = \inf\{u : X(u) \notin \mathbb{R}_+^N\}$ the exit time of $\mathbb{R}_+^N$. Then the distribution of $X(s \wedge \tau)$ has a singular component at the boundary where one of the coordinates is 0, corresponding to the absorbed mass, and a continuous density in the interior, which goes to zero at the boundaries. Let $p(t', x'|t, x)$ denote this continuous component of the distribution evaluated at $x' \in \mathbb{R}_+^N$ at time $t' \geq t$. For a fixed pair (t, x), it satisfies the forward Kolmogorov equation

$$\frac{\partial p}{\partial t'}(t', x'|t, x) = \mathcal{L}^{(N)*} p(t', x'|t, x), \tag{3.2}$$

with initial condition $p(t, x'|t, x) = \delta(x' - x)$, boundary condition $p(t', x'|t, x) = 0$ for $x'_1 \times \ldots \times x'_N = 0$, $t' \geq t$, and where $\mathcal{L}^{(N)*}$ is the adjoint operator of $\mathcal{L}^{(N)}$,

$$\mathcal{L}^{(N)*} = \frac{1}{2}\sum_{i,j=1}^{N} \rho_{ij} \frac{\partial^2}{\partial x'_i \partial x'_j} - \sum_{i=1}^{N} \mu_i \frac{\partial}{\partial x'_i}.$$

In PDE theory, the transition probability can be interpreted as Green's function, that is the fundamental solution to a suitable initial(-boundary) value problem. Thus, by time homogeneity, we use the notation $G(t' - t, x', x) = p(t', x'|t, x)$.

Once the Green's function has been constructed for a set of model parameters, the solution of (2.58)–(2.60) can be found as the solution of (3.2) by simple convolution of the Green's function with the appropriate inhomogeneous right-hand sides and boundary data. It is the strength of the Green's function concept that no further numerical solution of PDEs is required. We illustrate this in the following sections.

Applying the Green's theorem (Kythe (2011)) for (2.58)–(2.60), we get

$$V(t,x) = \int\limits_{\mathbb{R}_+^N} \psi(x')G(T-t,x',x)\,\mathrm{d}x' + \\ + \sum_{k=1}^{N} \int\limits_t^T \mathrm{d}t' \int\limits_{\mathbb{R}_+^{N-1}} \phi_{0,k}(t',x'_{-k})g_k(t'-t,x'_{-k},x)\,\mathrm{d}x'_{-k} - \int\limits_t^T \mathrm{d}t' \int\limits_{\mathbb{R}_+^N} \chi(t',x)G(t'-t,x',x)\,\mathrm{d}x, \tag{3.3}$$

where

$$g_k(t'-t,x'_{-k},x) = \frac{1}{2}\frac{\partial G}{\partial x'_k}(t'-t,x'_{0,k},x).$$

Thus, in order to solve the backward problem with non-homogeneous right-hand side and boundary conditions, it is sufficient to solve the forward problem for the Green's function with homogeneous right-hand side, boundary and terminal conditions.

3.3 Drift elimination

We can rewrite (3.2) in the vector form

$$\frac{\partial G}{\partial t} = \frac{1}{2}\nabla^* P \nabla G - \xi \cdot \nabla G, \tag{3.4}$$

where P is the correlation matrix.

Consider Cholesky decomposition of the matrix $P = L^*L$, where L is a lower-triangular matrix. Then, equation (3.4) can be rewritten in a matrix form

$$\frac{\partial G}{\partial t} = \frac{1}{2}(L\nabla)^*(L\nabla)G - \xi \cdot \nabla G, \tag{3.5}$$

Now consider the change of variables $y = L^{-1}x$. Then, the gradient transforms to $\nabla_x G(x) = L^{-1}\nabla_y G(y)$. Thus, in new variables we get rid of the cross-derivatives term

$$\frac{\partial G}{\partial t} = \frac{1}{2}\Delta_y G - \xi \cdot L^{-1}\nabla_y G, \tag{3.6}$$

The domain also changes to transformation of basis vectors V to $L^{-1}V$.

Now introduce

$$\tilde{G}(t'-t,y',y) = \exp\left(\xi \cdot L(y'-y) - \frac{1}{2}\xi^* P \xi (t'-t)\right) G(t'-t,y',y) \tag{3.7}$$

that satisfies the following equation

$$\frac{\partial \tilde{G}}{\partial t'} = \frac{1}{2}\Delta G, \tag{3.8}$$

with the same initial and boundary conditions.

Coming back to the original variables $x = Ly$ (and the original domain) we get the PDE without the drift term

$$\frac{\partial \tilde{G}}{\partial t} = \frac{1}{2}\sum_{i,j=1}^{d} \rho_{ij}\frac{\partial^2 \tilde{G}}{\partial x'_i \partial x'_j}, \tag{3.9}$$

As a result, we have showed how to eliminate the drift term. In the following, we consider (3.9) and ignore tilde for brevity.

3.4 One-dimensional problem

When $N = 1$, the problem (3.9) becomes

$$\frac{\partial G}{\partial t'}(t'-t,x',x) - \frac{1}{2}\frac{\partial^2 G}{\partial x'^2}(t'-t,x',x) = 0, \tag{3.10}$$

$$G(t'-t,x',0) = 0, \quad G(t'-t,x',x) \underset{x\ \to+\infty}{\longrightarrow} 0, \tag{3.11}$$

$$G(0,x',x) = \delta(x'-x). \tag{3.12}$$

Using the method of images, we represent $G(t'-t,x',x) = H(t'-t,x',x) - H(t'-t,-x',x)$, where $H(t'-t,y',y)$ is the solution of (3.10) and (3.12) on the infinite line. Then, both $H(t'-t,x',x)$ and $H(t'-t,-x',x)$ satisfy (3.10) as well as their difference. Moreover, it is easy to see that (3.11) and (3.12) are also satisfied.

Solving (3.10) with the initial condition (3.12) on $\mathbb{R}$, we get

$$H(t'-t,x',x) = \frac{1}{\sqrt{2\pi(t'-t)}} \mathrm{e}^{-\frac{(x'-x)^2}{2(t'-t)}}. \tag{3.13}$$

Therefore, we finally get

$$G(t'-t,x',x) = \frac{1}{\sqrt{2\pi(t'-t)}} \left(\mathrm{e}^{-\frac{(x'-x)^2}{2(t'-t)}} - \mathrm{e}^{-\frac{(x'+x)^2}{2(t'-t)}} \right). \tag{3.14}$$

3.5 Two-dimensional problem

When $N = 2$, the problem (3.9) becomes

$$\begin{cases} \dfrac{\partial G}{\partial t'} - \dfrac{1}{2}G_{x'x'} - \dfrac{1}{2}G_{y'y'} - \rho G_{x'y'} = 0, \\ G(t'-t,0,y',x,y) = 0, \quad G(t'-t,x',y',x,y) \underset{x'\ \to+\infty}{\longrightarrow} 0, \\ G(t'-t,x',0,x,y) = 0, \quad G(t'-t,x',y',x,y) \underset{y'\ \to+\infty}{\longrightarrow} 0, \\ G(0,x',y',x,y) = \delta(x'-x)\delta(y'-y). \end{cases} \tag{3.15}$$

First, we eliminate the cross–derivative by applying the change of variables

$$\begin{cases} \alpha(x,y) = x, \\ \beta(x,y) = -\dfrac{1}{\sqrt{1-\rho^2}}(\rho x - y). \end{cases} \tag{3.16}$$

It transforms the equation (3.15) to

$$\begin{cases} \dfrac{\partial G}{\partial t'} - \dfrac{1}{2}G_{\alpha'\alpha'} - \dfrac{1}{2}G_{\beta'\beta'} = 0, \\ G(0,\alpha',\beta',\alpha,\beta) = \delta(\alpha'-\alpha)\delta(\beta'-\beta), \\ \partial G(t'-t,\alpha',\beta',\alpha,\beta) = 0. \end{cases} \tag{3.17}$$

The domain has also changed from the positive quadrant to the interior of an angle characterized by $\cos(\bar{\omega}) = -\rho$.

The next step is to change the variables to polar coordinates to take advantage of the domain,

$$\begin{cases} \alpha = -r\sin(\varphi - \bar{\omega}), \\ \beta = r\cos(\varphi - \bar{\omega}). \end{cases} \tag{3.18}$$

Then, the equation (3.17) can be rewritten as

$$\begin{cases} \dfrac{\partial G}{\partial t'} - \dfrac{1}{2}\left(G_{r'r'} + \dfrac{1}{r'}G_{r'} + \dfrac{1}{r'^2}G_{\varphi'\varphi'}\right) = 0, \\ G(0, r', \varphi', r, \varphi) = \dfrac{1}{r}\delta(r' - r)\delta(\varphi' - \varphi), \\ G(t' - t, r', 0, r, \varphi) = 0, \quad G(t' - t, r', \bar{\omega}, r, \varphi) = 0, \quad G(r', \varphi', r, \varphi) \underset{r' \to +\infty}{\longrightarrow} 0. \end{cases} \tag{3.19}$$

To solve the equation (3.19), we use the separation of variables technique

$$G(t' - t, r', \varphi', r, \varphi) = g(t' - t, r', r)f(\varphi', \varphi). \tag{3.20}$$

Thus, we can rewrite the equation (3.19) as two separate equations

$$\begin{cases} g_{t'} = \dfrac{1}{2}\left(g_{r'r'} + \dfrac{1}{r'}g_{r'} + \dfrac{C}{r'^2}g\right), \\ g(0, r', r) = \dfrac{1}{r}\delta(r' - r), \\ g(t' - t, 0, r) = 0, \quad g(t' - t, r', r) \underset{r' \to +\infty}{\longrightarrow} 0, \end{cases} \tag{3.21}$$

and

$$\begin{cases} f_{\varphi'\varphi'} = Cf, \\ f(0, \varphi) = 0, \quad f(\bar{\omega}, \varphi) = 0. \end{cases} \tag{3.22}$$

Both equations can be solved analytically. It is also clear that $C < 0$. Thus, we denote $C = -\Lambda^2$. For equation (3.21) the solution is

$$g(\tau, r', r) = \frac{\mathrm{e}^{-\frac{r'^2 + r^2}{2\tau}}}{\tau} I_\Lambda\left(\frac{r'r}{\tau}\right), \tag{3.23}$$

where I_Λ is the modified Bessel function of the first kind defined as the solution of

$$\xi^2\frac{d^2I}{d\xi^2} + \xi\frac{dI}{d\xi} - (\xi^2 + \Lambda^2) = 0. \tag{3.24}$$

To verify the initial condition, we use the approximation of Bessel function $I_\Lambda(z)$ when $z \gg \Lambda$:

$$I_\Lambda(z) \approx \frac{\mathrm{e}^z}{\sqrt{2\pi z}}. \tag{3.25}$$

It yields

$$\lim_{\tau \to 0} g(\tau, r', r) = \lim_{\tau \to 0} \frac{\mathrm{e}^{-\frac{(r' - r)^2}{2\tau}}}{\sqrt{2\pi r' r \tau}} = \frac{1}{\sqrt{r'r}}\delta(r' - r) = \frac{1}{r'}\delta(r' - r). \tag{3.26}$$

For the equation (3.22), with respect to the boundary condition, we obtain $\Lambda_n = \frac{\pi n}{\bar{\omega}}$ and the solution is

$$f_n(\varphi', \varphi) = A_n \sin\left(\frac{n\pi\varphi'}{\bar{\omega}}\right). \tag{3.27}$$

Once we have solved the equations and found the corresponding eigenvalues and eigenfunctions, we can write the eigenvalue expansion of the Green's function

$$G(\tau, r', \varphi', r, \varphi) = \frac{\mathrm{e}^{-\frac{r'^2+r^2}{2\tau}}}{\tau} \sum_{n=1}^{\infty} C_n I_{\frac{n\pi}{\bar{\omega}}} \left(\frac{r'r}{\tau}\right) \sin\left(\frac{n\pi\varphi'}{\bar{\omega}}\right). \tag{3.28}$$

The coefficients C_n can be computed by applying the initial condition. Using the fact that

$$g(\tau, r', r) \underset{\tau \to 0}{\longrightarrow} \frac{1}{r}\delta(r' - r),$$

the initial condition yields

$$\sum_{n=1}^{\infty} C_n \sin\left(\frac{n\pi\varphi'}{\bar{\omega}}\right) = \delta(\varphi' - \varphi).$$

Multiplying both parts by $\sin\left(\frac{m\pi\varphi'}{\bar{\omega}}\right)$ and integrating from 0 to $\bar{\omega}$, we find the coefficients $C_n = \frac{2}{\bar{\omega}} \sin\left(\frac{n\pi\varphi}{\bar{\omega}}\right)$. Thus, we finally get

$$G(\tau, r', \varphi', r, \varphi) = \frac{2\mathrm{e}^{-\frac{r'^2+r^2}{2\tau}}}{\bar{\omega}\tau} \sum_{n=1}^{\infty} I_{\frac{n\pi}{\bar{\omega}}} \left(\frac{r'r}{\tau}\right) \sin\left(\frac{n\pi\varphi'}{\bar{\omega}}\right) \sin\left(\frac{n\pi\varphi}{\bar{\omega}}\right). \tag{3.29}$$

3.6 Three-dimensional problem

The three-dimensional problem is much more difficult than the one- and two-dimensional ones. For several particular cases of the correlation matrix, Escobar et al. (2013) offered solutions to the three-dimensional problem by the method of images. Unfortunately, it cannot be extended to general correlation matrices. A semi-analytical solution was proposed in Lipton and Savescu (2014), where a separation of variables technique is used and the problem is split into radial and angular PDEs. The radial PDE is solved analytically, while the angular PDE is solved by finite elements. In this section, we use the separation of variables technique and solve the angular PDE semi-analytically, and thus provide a "fully semi-analytical" representation of the Green's function. By this we mean that the solution is expressed as a (double) infinite series of terms involving special functions (hypergeometric and Bessel). We also require the solution of a countably infinite-dimensional non-linear eigenvalue problem. In practice, the sums can be truncated to a relatively small number of terms (low tens) without significant loss of accuracy, as evidenced by our numerical tests, and similarly for the dimension of the non-linear eigenvalue problem.

3.6.1 Problem formulation and preliminaries

When $N = 3$, the problem (3.9) becomes

$$\frac{\partial G}{\partial t'} = \frac{1}{2}G_{x'x'} + \frac{1}{2}G_{y'y'} + \frac{1}{2}G_{z'z'} + \rho_{xy}G_{x'y'} + \rho_{xz}G_{x'z'} + \rho_{yz}G_{y'z'}, \tag{3.30}$$

with the boundary and initial conditions

$$\begin{cases} G(0,x',y',z',x,y,z) = \delta(x'-x)\delta(y'-y)\delta(z'-z), & (x',y',z') \in \mathbb{R}^3_+, \\ G(t'-t,0,y',z',x,y,z) = 0, \quad G(t'-t,x',y',z',x,y,z) \underset{x' \to+\infty}{\longrightarrow} 0, & (t',y',z') \in \mathbb{R}^3_+, \\ G(t'-t,x',0,z',x,y,z) = 0, \quad G(t'-t,x',y',z',x,y,z) \underset{y' \to+\infty}{\longrightarrow} 0, & (t',x',z') \in \mathbb{R}^3_+, \\ G(t'-t,x',y',0,x,y,z) = 0, \quad G(t'-t,x',y',z',x,y,z) \underset{z' \to+\infty}{\longrightarrow} 0, & (t',x',y') \in \mathbb{R}^3_+, \end{cases}$$

for given $(t,x,y,z) \in \mathbb{R}^4_+$.

To eliminate the cross-derivatives, we introduce the following change of variables (Lipton et al. (2014), Lipton and Savescu (2014))

$$\begin{cases} \alpha(x,y,z) = x, \\ \beta(x,y,z) = -\dfrac{1}{\bar\rho_{xy}}(\rho_{xy}x - y), \\ \gamma(x,y,z) = \dfrac{1}{\bar\rho_{xy}\chi}\left[(\rho_{xy}\rho_{yz} - \rho_{xz})x + (\rho_{xy}\rho_{xz} - \rho_{yz})y + \bar\rho^2_{xy}z\right], \end{cases}$$

where $\bar\rho_{xy} = \sqrt{1-\rho^2_{xy}}$ and $\chi = \sqrt{1-\rho^2_{xy}-\rho^2_{xz}-\rho^2_{yz}+2\rho_{xy}\rho_{xz}\rho_{yz}}$, which is the Cholesky decomposition, we used earlier. It transforms the equation (3.30) to

$$\begin{cases} \dfrac{\partial G}{\partial t'} = \dfrac{1}{2}G_{\alpha'\alpha'} + \dfrac{1}{2}G_{\beta'\beta'} + \dfrac{1}{2}G_{\gamma'\gamma'}, \\ G(0,\alpha',\beta',\gamma',\alpha,\beta,\gamma) = \delta(\alpha'-\alpha)\delta(\beta'-\beta)\delta(\gamma'-\gamma), \end{cases} \tag{3.31}$$

with zero boundary conditions on the new domain

$$D = \{\omega_1 e_1 + \omega_2 e_2 + \omega_3 e_3 | \omega_i > 0\},$$

where

$$e_1 = (0,0,1), \quad e_2 = \left(0, \frac{\chi}{\bar\rho_{xy}\bar\rho_{xz}}, -\frac{\rho_{yz}-\rho_{xz}\rho_{xy}}{\bar\rho_{xy}\bar\rho_{xz}}\right), \quad e_3 = \left(\frac{\chi}{\bar\rho_{yz}}, -\frac{\rho_{xy}\chi}{\bar\rho_{xy}\bar\rho_{yz}}, -\frac{\rho_{xz}-\rho_{yz}\rho_{xy}}{\bar\rho_{xy}\bar\rho_{yz}}\right),$$

Finally, we rewrite the equation in spherical coordinates (r,φ,θ) to take advantage of the wedge-shape of the domain, where $r \in (0,\infty)$, $\varphi \in (0,\bar\omega)$ with $\bar\omega = \arccos(-\rho_{xy})$, and $\theta \in (0,\Theta(\varphi))$, where $\Theta(\varphi)$ can be implicitly defined as

$$\varphi(\omega) = \arccos\left(\frac{1-\rho_{xy}\omega}{\sqrt{1-2\rho_{xy}\omega+\omega^2}}\right), \tag{3.32}$$

$$\Theta(\omega) = \arccos\left(-\frac{\rho_{yz}-\rho_{xz}\rho_{xy}+\omega(\rho_{xz}-\rho_{yz}\rho_{xy})}{\sqrt{\bar\rho_{xy}(\bar\rho^2_{xz}-2\omega(\rho_{xy}-\rho_{xz}\rho_{yz})+\omega^2\bar\rho^2_{yz})}}\right), \tag{3.33}$$

where $\bar\rho = \sqrt{1-\rho^2}$ for $\rho \in \{\rho_{xy},\rho_{yz}\}$.

The final initial-boundary value problem for the Green's function has the form

$$\begin{cases} \dfrac{\partial G}{\partial t'} - \dfrac{1}{2}\left[\dfrac{1}{r'}\dfrac{\partial^2}{\partial r'^2}(r'G) + \dfrac{1}{r'^2}\left(\dfrac{1}{\sin^2\theta'}G_{\varphi'\varphi'} + \dfrac{1}{\sin\theta'}\dfrac{\partial}{\partial\theta'}(\sin\theta' G_{\theta'}\right)\right] = 0, \\ G(0,r',\varphi',\theta',r,\varphi,\theta) = \dfrac{1}{r^2\sin\theta}\delta(r'-r)\delta(\varphi'-\varphi)\delta(\theta'-\theta), \\ G(t'-t,r',0,\theta',r,\varphi,\theta) = 0, \qquad G(t'-t,r',\bar{\omega},\theta',r,\varphi,\theta) = 0, \\ G(t'-t,r',\varphi',0,r,\varphi,\theta) = 0, \qquad G(t'-t,r',\varphi',\Theta(\varphi')\mathrm{e}'r,\varphi,\theta) = 0, \\ G(t'-t,0,\varphi',\theta',r,\varphi,\theta) = 0, \qquad G(t'-t,r',\varphi',\theta',r,\varphi,\theta) \underset{r'\to+\infty}{\longrightarrow} 0. \end{cases} \tag{3.34}$$

3.6.2 Semi-analytical solution by expansion

In this section, we show how the solution to (3.34) can be given via an eigenvalue expansion. First, we apply separation of variables

$$G(t,r',\varphi',\theta',r,\varphi,\theta) = g(t,r',r)\Psi(\varphi',\theta',\varphi,\theta),$$

then (3.34) decomposes into two separate (initial-)boundary value problems

$$\begin{cases} g_{t'} = \dfrac{1}{2}\left(\dfrac{1}{r'}\dfrac{\partial^2}{\partial r'^2}(r'g) - \dfrac{\Lambda^2}{r'^2}g\right), \\ g(0,r',r) = \dfrac{1}{r}\delta(r'-r), \\ g(t'-t,0,r) = 0, \quad g(t'-t,r',r) \underset{r'\to+\infty}{\longrightarrow} 0, \end{cases} \tag{3.35}$$

and

$$\begin{cases} \dfrac{1}{\sin^2\theta'}\Psi_{\varphi'\varphi'} + \dfrac{1}{\sin\theta'}\dfrac{\partial}{\partial\theta'}(\sin\theta'\Psi_{\theta'}) = -\Lambda^2\Psi, \\ \Psi(0,\theta',\varphi,\theta) = 0, \Psi(\bar{\omega},\theta',\varphi,\theta) = 0, \Psi(\varphi',0,\varphi,\theta) = 0, \Psi(\varphi',\Theta(\varphi'),\varphi,\theta) = 0, \end{cases} \tag{3.36}$$

where $\Lambda^2 \in \mathbb{R}_+$ is an eigenvalue. The eigenvalue is real and positive because the differential operator on the left-hand side of (3.36) is symmetric and positive definite with respect to a certain weighted inner product, as is seen from the variational formulation given in equation (40) in Lipton and Savescu (2014).

We note an asymmetry between (3.35), which contains the time variable and hence an initial condition is imposed, and (3.36), which is stationary and hence does not match any initial condition. The initial condition of the solution in (3.34) will eventually be ensured by superposition in Section 3.6.5. Since (3.36) does not depend on (φ,θ) here (due to the ignorance of the initial condition), we further write $\Psi(\varphi',\theta',\varphi,\theta) = \Psi(\varphi',\theta')$ for simplicity.

The PDE (3.35) can be solved analytically, with solution

$$g(t,r',r) = \frac{\mathrm{e}^{-\frac{r^2+r'^2}{2t}}}{t\sqrt{rr'}} I_{\sqrt{\Lambda^2+\frac{1}{4}}}\left(\frac{rr'}{t}\right),$$

where $I_a(b)$ is the modified Bessel function of the second kind.

To verify the initial condition is satisfied, we again use the approximation (3.25):

$$\lim_{t\to 0} g(t,r',r) = \lim_{t\to 0} \frac{\mathrm{e}^{-\frac{(r'-r)^2}{2t}}}{\sqrt{2\pi r' r t}} = \frac{1}{\sqrt{r'r}}\delta(r'-r) = \frac{1}{r'}\delta(r'-r).$$

The eigenvalue problem for the angular PDE (3.36) was solved numerically in Lipton and Savescu (2014) by a finite element method. In the next section we demonstrate this method.

3.6.3 Finite element method for angular PDE

Lipton and Savescu (2014) solved the equation (3.36) numerically via the finite element method. In this section we briefly explain the method.

It can be shown that there is an infinite sequence of eigenvalues $0 \le \Lambda_1 \le \Lambda_2 \le \dots$ (see Lipton and Savescu (2014)).

We reformulate the problem (3.36) in a weak formulation. We use the test functions Ψ' and integrate over the domain,

$$\iint\limits_{\Omega} \frac{1}{\sin\theta'}\Psi_{\varphi'}\Psi'_{\varphi'}\,\mathrm{d}\omega + \iint\limits_{\Omega} \sin\theta'\Psi_{\theta'}\Psi'_{\theta'}\,\mathrm{d}\omega = \Lambda^2 \iint\limits_{\Omega} \Psi\Psi'\sin\theta'\,\mathrm{d}\omega. \tag{3.37}$$

For this non-rectangular shape of the domain we can construct a triangular mesh and look for the finite-dimensional solution on this mesh. See mesh construction section for details.

The dimension n is the number of free points on the mesh. Denote by $\Phi_i, i = 1,\dots,n$ the basis of this space.

We consider linear basis functions on each triangle. Hence, there are only three independent (non-zero) basis functions on each triangle T. We denote by (φ'_1,θ'_1), (φ'_2,θ'_2), and (φ'_3,θ'_3) the vertices of triangle T and by Φ_1,Φ_2 and Φ_3 the corresponding basis functions that are defined by

$$\Phi_i(\varphi'_j,\theta'_j) = \begin{cases} 1, i = j, \\ 0, i \ne j, \end{cases}$$

and can be represented as $\Phi_i(\varphi',\theta') = a_i + b_i\varphi' + c_i\theta', (\varphi',\theta') \in T, i = 1,2,3$, where the coefficients a_i, b_i, c_i can be found by solving 3×3 system of linear equations.

Thus, the solution of finite-dimensional problem can be represented as $\Psi(\varphi',\theta') = \sum_{i=1}^{n}\Phi_i(\varphi',\theta')\psi_i$ and equation (3.37) can be written as

$$K\Psi = \Lambda^2 M\Psi, \tag{3.38}$$

where K and M are matrices with elements

$$K_{ij} = \iint\limits_{\Omega} (A\nabla\Phi_j)\cdot\nabla\Phi_i\,\mathrm{d}\omega,$$

$$M_{ij} = \iint\limits_{\Omega} \sin\theta'\Phi_i\Phi_j\,\mathrm{d}\omega,$$

with matrix A

$$A = \begin{pmatrix} \frac{1}{\sin\theta'} & 0 \\ 0 & \sin\theta' \end{pmatrix}.$$

Denote by N_t the number of triangles. Then, equations for K_{ij} and M_{ij} can be further simplified to

$$K_{ij} = \sum_{k=1}^{N_t} \int_{T_k} \frac{1}{\sin\theta'} \frac{\partial\Phi_i}{\partial\varphi'} \frac{\partial\Phi_j}{\partial\varphi'} \,\mathrm{d}\omega + \sum_{k=1}^{N_t} \int_{T_k} \sin\theta' \frac{\partial\Phi_i}{\partial\theta'} \frac{\partial\Phi_j}{\partial\theta'} \,\mathrm{d}\omega,$$
$$M_{ij} = \sum_{k=1}^{N_t} \int_{T_k} \sin\theta' \Phi_i \Phi_j \,\mathrm{d}\omega.$$

In order to compute integrals over T_k we approximate them by the quadrature rule

$$\int_{T_k} f(\varphi', \theta') \,\mathrm{d}\omega = \mathrm{Area}_{T_k} f(\bar{\varphi}, \bar{\theta}),$$

where $(\bar{\varphi}, \bar{\theta})$ is the centroid of the triangle T_k.

Taking into account that over triangles the basis functions are linear and, therefore, their derivatives are constant, the computation of K_{ij} can be further simplified.

To solve the equation (3.38), we perform a Cholesky decomposition of the symmetric matrix $M = \mathcal{M}\mathcal{M}^T$, then

$$\mathcal{M}^{-1} K \Psi = \Lambda^2 \mathcal{M}^T \Psi.$$

We introduce the matrix $C = \mathcal{M}^{-1} K (\mathcal{M}^{-1})^T$, then the last equation can be rewritten as

$$C(\mathcal{M}^T)\Psi = \Lambda^2(\mathcal{M}^T\Psi). \tag{3.39}$$

The equation (3.39) is the matrix eigenvalue problem for matrix C and can be solved via standard numerical methods. Once we solved the problem (3.39), the eigenvalues of the original problem can be found as $\mathcal{M}^{-1}E$, where E is an eigenvector of (3.39). As a result, we have found the first N eigenvalues and the corresponding eigenvectors.

3.6.3.1 Mesh construction

In this section, we briefly describe the triangular mesh construction algorithm. Lipton and Savescu (2014) presented an iterative algorithm based on ideas of Persson (2004), where the nodes of the mesh are adjusted in each iteration based on current element sizes. See Algorithm 1 for the algorithm description.

The simplest case is to choose a uniform distribution of the edge size. But, on our domain, there is a singularity on the boundary, therefore we choose a size distribution that selects smaller elements near the boundary and larger in the center.

An example of the mesh is in Figure 3.1.

Algorithm 1 Algorithm for constructing an adaptive mesh

Require: $X_1.X_2, Y_1, Y_2$ — boundary box of the domain
Require: $d(x, y)$ — distance function to the closest boundary (negative inside the domain)
Require: $h(x, y)$ — element size function (gives the relative element size distribution over the domain)
1: Build a mesh with equally spaced points for the bounding box of the domain
2: Remove points outside the domain
3: Rejection method: reject points inside the domain with probabilities proportional $1/h(x, y)^2$
4: **while** i $<$ MAXITER **do**
5: Delaunay triangulation using the divide and conquer algorithm described in Guibas and Stolfi (1985)
6: Assemble triangles obtained through the Delaunay procedure
7: **for** each triangle **do**
8: Compute centroid
9: If centroid is outside the domain, remove triangle from the list
10: **end for**
11: Move mesh points based on current edge lengths using ideas described in Persson (2004)
12: Bring points that have moved outside of the domain back to the boundary
13: i = i + 1
14: **end while**

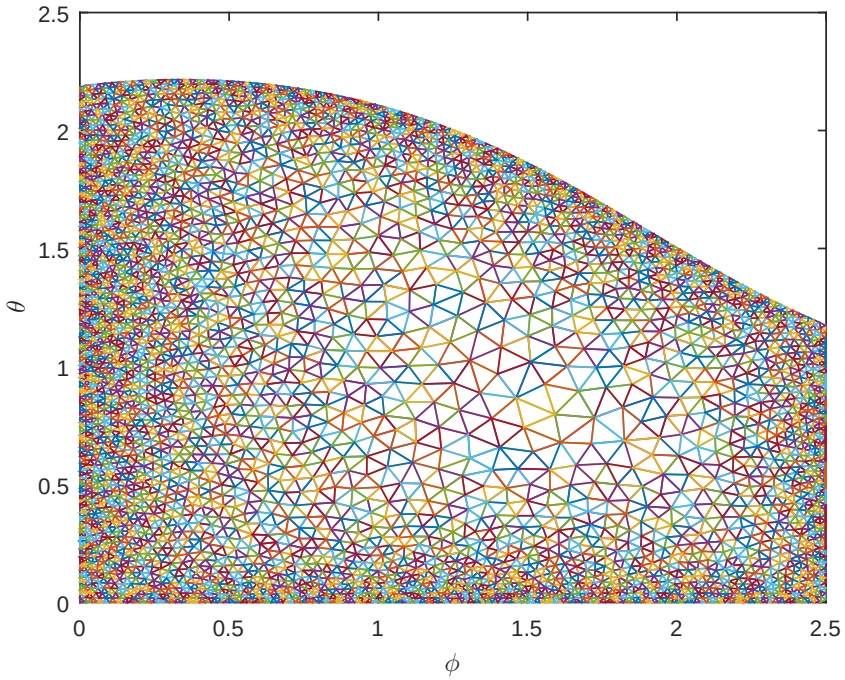


Figure 3.1: An example of mesh for the domain obtained for $\rho_{xy} = 0.8, \rho_{xz} = 0.2, \rho_{yz} = 0.5$ for 4000 points after 100 iterations of the algorithm (the color have no significance).

3.6.4 Semi-analytical method for angular PDE

In this section, we consider an alternative, semi-analytical approach.

We begin by further simplifying the problem by the extra transformation (as in Lipton et al. (2014) for the stationary case)

$$\zeta' = \ln \tan \theta'/2, \tag{3.40}$$

which changes the domain to the semi-infinite strip $-\infty < \zeta \leq Z(\varphi) = \ln \tan \Theta(\varphi)/2$ and the eigenvalue problem to

$$\begin{cases} \Psi_{\varphi'\varphi'} + \Psi_{\zeta'\zeta'} = -4\Lambda^2 \dfrac{\mathrm{e}^{2\zeta'}}{(1+\mathrm{e}^{2\zeta'})^2}\Psi, \\ \Psi(0,\zeta') = \Psi(\bar{\omega},\zeta') = \lim\limits_{\zeta'\to-\infty} \Psi(\varphi',\zeta') = \Psi(\varphi', Z(\varphi')) = 0. \end{cases} \tag{3.41}$$

Denote $\lambda = \frac{1}{2} + \sqrt{\frac{1}{4} + \Lambda^2}$. Then, (3.41) becomes

$$\Psi_{\varphi'\varphi'} + \Psi_{\zeta'\zeta'} = -\frac{\lambda(\lambda-1)}{\cosh^2(\zeta')}\Psi.$$

We shall find solutions in the form $\Psi(\varphi',\zeta') = \Phi(\varphi') \cdot \Upsilon(\zeta')$. Thereby, we can split (3.41) into the system

$$\Phi'' = -k^2\Phi, \tag{3.42}$$

$$\Upsilon'' = \left(k^2 - \frac{\lambda(\lambda-1)}{\cosh^2(\zeta')}\right)\Upsilon, \tag{3.43}$$

where $k^2 \in \mathbb{R}_+$ is an eigenvalue and with boundary conditions

$$\Phi(0) = 0, \quad \Phi(\bar{\omega}) = 0, \quad \Upsilon(\zeta') \underset{\zeta \to -\infty}{\longrightarrow} 0, \quad \Psi(\varphi', Z(\varphi')) = 0. \tag{3.44}$$

Equation (3.42) with the first and second boundary conditions in (3.44) can be solved as

$$\Phi_n(\varphi') = c_n \sin(k_n \varphi'),$$

where $k_n = \pi n / \bar{\omega}$.

Equation (3.43) is the Schrödinger equation with Pöschl–Teller potential. In the next section we show how to solve it.

3.6.4.1 Schrödinger equation with Pöschl–Teller potential

Consider (3.43) with boundary condition

$$\Upsilon(-\infty) = 0. \tag{3.45}$$

Fundamental solution We follow Flügge (2012) for the solution of (3.43). Consider the change of variables $y = \cosh^2(\zeta)$. Then, (3.43) can be rewritten as

$$y(1-y)\Upsilon'' + \left(\frac{1}{2} - y\right)\Upsilon' + \left(\frac{k^2}{4} - \frac{\lambda(\lambda-1)}{4y}\right)\Upsilon = 0.$$

Then, if we consider $v(y) = y^{-\lambda/2}\Upsilon(y)$, we get the hypergeometric equation

$$y(1-y)v'' + \left(\left(\lambda + \frac{1}{2}\right) - (\lambda+1)y\right)v' + \frac{1}{4}\left(k^2 - \lambda^2\right)v = 0.$$

The solution of this equation is

$$v(y) = A \cdot {}_2F_1\left(a, b, \frac{1}{2}, 1-y\right) + B\sqrt{1-y} \cdot {}_2F_1\left(a + \frac{1}{2}, b + \frac{1}{2}, \frac{3}{2}, 1-y\right),$$

where

$$a = \frac{1}{2}(\lambda - k), \quad b = \frac{1}{2}(\lambda + k).$$

As a fundamental system we consider standard odd and even real solutions:

$$\Upsilon_e(\zeta) = \cosh^\lambda(\zeta) {}_2F_1\left(a, b, \frac{1}{2}, -\sinh^2(\zeta)\right),$$
$$\Upsilon_o(\zeta) = \cosh^\lambda(\zeta)\sinh(\zeta) {}_2F_1\left(a + \frac{1}{2}, b + \frac{1}{2}, \frac{3}{2}, -\sinh^2(\zeta)\right),$$

where ${}_2F_1(a, b, c, z)$ is the Gaussian hypergeometric function.

Any solution can be represented as a linear combination of the fundamental system,

$$\Upsilon(\zeta) = A\Upsilon_e(\zeta) + B\Upsilon_o(\zeta). \tag{3.46}$$

Now consider an asymptotic expansion for large ζ. Unlike Flügge (2012), who consider the asymptotics at $\pm\infty$ (see the case $Z \equiv \infty$ below), if $Z(\varphi)$ is bounded above, we only consider the asymptotics at $-\infty$.

General boundary conditions From the known expansion of the hypergeometric function we get

$$\Upsilon_e(\zeta) \to 2^{-\lambda} \mathrm{e}^{\lambda|\zeta|} \Gamma\left(\frac{1}{2}\right) \left[\frac{\Gamma(b-a)}{\Gamma(b)\Gamma\left(\frac{1}{2}-a\right)} 2^{2a} \mathrm{e}^{-2a|\zeta|} + \frac{\Gamma(a-b)}{\Gamma(a)\Gamma\left(\frac{1}{2}-b\right)} 2^{2b} \mathrm{e}^{-2b|\zeta|} \right]$$

and

$$\Upsilon_o(\zeta) \to -2^{-\lambda-1} \mathrm{e}^{(\lambda+1)|\zeta|} \Gamma\left(\frac{3}{2}\right) \left[\frac{\Gamma(b-a)}{\Gamma\left(b+\frac{1}{2}\right)\Gamma(1-a)} 2^{2a+1} \mathrm{e}^{-(2a+1)|\zeta|} + \right.$$
$$\left. + \frac{\Gamma(a-b)}{\Gamma\left(a+\frac{1}{2}\right)\Gamma(1-b)} 2^{2b+1} \mathrm{e}^{-(2b+1)|\zeta|} \right].$$

Assume $k > 0$. First, consider the case $\lambda - k = 2l$. Thus, $\Gamma\left(\frac{1}{2}-a\right)$ is singular and $\Upsilon_e(\zeta)$ satisfies the condition (3.45). Similar, if $\lambda - k = 2l+1$, then $\Gamma(1-a)$ is singular and $\Upsilon_o(\zeta)$ satisfies the condition (3.45).

Hence, in order to satisfy the condition (3.45) we can choose

$$A = \begin{cases} \Gamma(b)\Gamma\left(\frac{1}{2}-a\right), & \lambda - k \notin \mathbb{Z}_+, \\ 1, & \lambda - k = 2l \\ 0, & \lambda - k = 2l+1 \end{cases}, B = \begin{cases} 2\Gamma\left(b+\frac{1}{2}\right)\Gamma(1-a), & \lambda - k \notin \mathbb{Z}_+, \\ 0, & \lambda - k = 2l, \\ 1, & \lambda - k = 2l+1. \end{cases}$$

Using identities for hypergeometric functions we can simplify (3.46) to

$$\Upsilon(\zeta) = c \cdot \mathrm{e}^{k\zeta} {}_2F_1\left(\lambda, 1-\lambda, 1+k, \frac{\mathrm{e}^{2\zeta}}{1+\mathrm{e}^{2\zeta}}\right),$$

for all λ and k.

Special case $\rho_{xy} = \rho_{xz} = 0$, $\rho_{yz} = 1$, i.e. $Z \equiv \infty$ Consider first the case when $Z(\varphi)$ is unbounded. Now we have to take into account also the asymptotics at $+\infty$. We note that $\Upsilon_e(-\zeta) = \Upsilon_e(\zeta)$ and $\Upsilon_o(-\zeta) = -\Upsilon_o(\zeta)$. The first term in $\Upsilon_o(\zeta)$ and $\Upsilon_e(\zeta)$ goes to infinity unless $\Gamma(\frac{1}{2}-a)$ and $\Gamma(1-a)$ are singular, respectively. The gamma function is singular if the argument is a negative integer. Thus,

$$\lambda_l^r = \frac{\pi r}{\bar{\omega}} + l$$

and

$$\Psi_l^r(\varphi, \zeta) = \begin{cases} c \sin(k_r \varphi) \Upsilon_e(\zeta), l \text{ is even}, \\ c \sin(k_r \varphi) \Upsilon_o(\zeta), l \text{ is odd}. \end{cases}$$

Special case $\rho_{xy} = \rho_{xz} = \rho_{yz} = 0$, i.e. $Z \equiv 0$ In this case $\bar{\omega} = \frac{\pi}{2}$. We can easily see that $\Upsilon_e(0) = 1$ and $\Upsilon_o(0) = 0$ for any k and λ. Thus, $A = 0, B = 1$. Now we want to find the corresponding eigenvalues. Applying boundary condition (3.45), we get

$$\lambda_l = 2l + 1, n \geq 0,$$

and the corresponding eigenvectors are

$$\Psi_l^r(\varphi, \zeta) = c \sin(k_r \varphi) \Upsilon_o(\zeta),$$

where $1 \leq r \leq l$.

3.6.4.2 Expansion for $\Psi(\varphi', \zeta')$

As we showed in the previous section

$$\Upsilon_n(\zeta') = c_n \mathrm{e}^{k_n \zeta'} {}_2F_1\left(\lambda, 1-\lambda, 1+k_n, \frac{\mathrm{e}^{2\zeta'}}{1+\mathrm{e}^{2\zeta'}}\right),$$

where ${}_2F_1(a, b, c, z)$ is the Gaussian hypergeometric function. Thus, we find $\Psi(\varphi', \zeta')$ in the form

$$\Psi(\varphi', \zeta') = \sum_{n=1}^{\infty} c_n \sin(k_n \varphi') \mathrm{e}^{k_n \zeta'} {}_2F_1\left(\lambda, 1-\lambda, 1+k_n, \frac{\mathrm{e}^{2\zeta'}}{1+\mathrm{e}^{2\zeta'}}\right). \tag{3.47}$$

As an aside, we note that ${}_2F_1\left(\lambda, 1-\lambda, 1+k_n, \frac{\mathrm{e}^{2\zeta'}}{1+\mathrm{e}^{2\zeta'}}\right) = 1$ for $\lambda = 0$, which gives

$$\Psi_0(\varphi', \zeta') = \sum_{n=1}^{\infty} c_n \sin(k_n \varphi') \mathrm{e}^{k_n \zeta'}$$

and coincides with the expression in Lipton et al. (2014) for the stationary problem.

To find the coefficients c_n and eigenvalues λ, we use the fourth condition in (3.44),

$$\sum_{n=1}^{\infty} c_n \sin(k_n \varphi) \mathrm{e}^{k_n Z(\varphi')} {}_2F_1\left(\lambda, 1-\lambda, 1+k_n, \frac{\mathrm{e}^{2Z(\varphi')}}{1+\mathrm{e}^{2Z(\varphi')}}\right) = 0 \tag{3.48}$$

for all $\varphi' \in [0, \bar{\omega}]$.

We introduce the integrals

$$\begin{aligned} T_{mn}(\lambda) &= \langle f_m, f_n \rangle = \int_0^{\bar{\omega}} f_m(\varphi') f_n(\varphi') \, \mathrm{d}\varphi', \\ f_n(\varphi') &= \sin(k_n \varphi') \mathrm{e}^{k_n Z(\varphi')} {}_2F_1\left(\lambda, 1-\lambda, 1+k_n, \frac{\mathrm{e}^{2Z(\varphi')}}{1+\mathrm{e}^{2Z(\varphi')}}\right), \end{aligned} \tag{3.49}$$

and note that T_{mn} is a function of λ only. Then, applying (3.48), we get

$$\sum_{n=1}^{\infty} T_{mn}(\lambda) c_n = 0, \quad \forall m \geq 1. \tag{3.50}$$

We defer the solution of (3.50) for λ and $(c_n)_{n \geq 1}$ to Section 3.6.6.

3.6.5 Final expression for Green's function

Now we can write the expression for the Green's function using the eigenvalue expansion

$$\begin{aligned} G(\tau, r', \varphi', \zeta', r, \varphi, \zeta) &= \sum_{l=1}^{\infty} a_l g_l(\tau, r') \Psi_l(\varphi', \zeta') \\ &= \sum_{l=1}^{\infty} a_l g_l(\tau, r') \sum_{n=1}^{\infty} c_n^l \Phi_n(\varphi') \Upsilon_n^l(\zeta') \\ &= \frac{\mathrm{e}^{-\frac{r'^2+r^2}{2\tau}}}{\tau\sqrt{r'r}} \sum_{l=1}^{\infty} a_l I_{\lambda_l}\left(\frac{r'r}{\tau}\right) \times \sum_{n=1}^{\infty} c_n^l \sin(k_n \varphi') \mathrm{e}^{k_n \zeta'} {}_2F_1\left(\lambda_l, 1-\lambda_l, 1+k_n, \frac{\mathrm{e}^{2\zeta'}}{1+\mathrm{e}^{2\zeta'}}\right) \end{aligned} \tag{3.51}$$

where the (countably many) eigenvalues λ_l and coefficients c_n^l can be determined as the solution of the nonlinear eigenvalue problem (3.50) and $k_n = \frac{\pi n}{\omega}$.

The coefficients a_l can be determined by imposing the initial condition

$$G(0, r', \varphi', \theta', r, \varphi, \theta) = \frac{1}{r^2 \sin\theta} \delta(r' - r)\delta(\varphi' - \varphi)\delta(\theta' - \theta).$$

We have already ensured the initial condition for the radial part, thus we need

$$\sum_{l=1}^{\infty} a_l \Psi_l(\varphi', \theta') = \frac{1}{\sin\theta} \delta(\varphi' - \varphi)\delta(\theta' - \theta). \tag{3.52}$$

From the weak formulation of the angular PDE it is easy to see that the eigenvectors are orthogonal in the scalar product weighted by $\sin\theta'$ (Lipton and Savescu (2014)). Multiplying (3.52) by $\Psi_m(\varphi', \theta') \sin\theta'$ and integrating over the whole domain,

$$a_l = \iint\limits_{\Omega} \frac{1}{\sin\theta} \Psi_l(\varphi', \theta') \sin\theta' \delta(\varphi' - \varphi)\delta(\theta' - \theta)\, \mathrm{d}\varphi' \mathrm{d}\theta' = \Psi_l(\varphi, \theta).$$

Substituting the expression for a_l from the last equation into (3.51), expanding the expression for $\Psi_l(\varphi', \zeta')$ and $\Psi_l(\varphi, \zeta)$, we get

$$\begin{aligned} G(\tau, r', \varphi', \zeta', r, \varphi, \zeta) = {} & \frac{\mathrm{e}^{-\frac{r'^2 + r^2}{2\tau}}}{\tau\sqrt{r'r}} \sum_{l=1}^{\infty} I_{\lambda_l}\left(\frac{r'r}{\tau}\right) \times \\ & \times \left(\sum_{n=1}^{\infty} c_n^l \sin(k_n \varphi') \mathrm{e}^{k_n \zeta'} {}_2F_1\left(\lambda_l, 1 - \lambda_l, 1 + k_n, \frac{\mathrm{e}^{2\zeta'}}{1 + \mathrm{e}^{2\zeta'}}\right)\right) \times \\ & \times \left(\sum_{n=1}^{\infty} c_n^l \sin(k_n \varphi) \mathrm{e}^{k_n \zeta} {}_2F_1\left(\lambda_l, 1 - \lambda_l, 1 + k_n, \frac{\mathrm{e}^{2\zeta}}{1 + \mathrm{e}^{2\zeta}}\right)\right). \end{aligned}$$

Returning to the variable θ, we have

$$\begin{aligned} G(\tau, r', \varphi', \theta', r, \varphi, \theta) = {} & \frac{\mathrm{e}^{-\frac{r'^2 + r^2}{2\tau}}}{\tau\sqrt{r'r}} \sum_{l=1}^{\infty} I_{\lambda_l}\left(\frac{r'r}{\tau}\right) \times \\ & \times \left(\sum_{n=1}^{\infty} c_n^l \sin(k_n \varphi') \tan^{k_n}\left(\frac{\theta'}{2}\right) {}_2F_1\left(\lambda_l, 1 - \lambda_l, 1 + k_n, \sin^2\frac{\theta'}{2}\right)\right) \times \\ & \times \left(\sum_{n=1}^{\infty} c_n^l \sin(k_n \varphi) \tan^{k_n}\left(\frac{\theta}{2}\right) {}_2F_1\left(\lambda_l, 1 - \lambda_l, 1 + k_n, \sin^2\frac{\theta}{2}\right)\right). \end{aligned} \tag{3.53}$$

3.6.6 Solution of the nonlinear eigenvalue problem

In this section, we first give properties of the eigenvalue problem (3.50). Then, we describe several methods to solve nonlinear eigenvalue problems and outline how the methods are applied in the tests.

3.6.6.1 Truncation and the smallest (linear) eigenvalue of $T(\lambda)$

If we truncate the sum (3.50) after N terms, the resulting equation can be rewritten in matrix form

$$T(\lambda) \cdot c = 0, \tag{3.54}$$

where $T(\lambda) \in \mathbb{R}^{N \times N}, c \in \mathbb{R}^N$.

Equation (3.54) is a nonlinear eigenvalue problem for a symmetric positive semi-definite matrix (as the matrix T is a Gram matrix). Eigenvalues λ of this truncated problem have to satisfy the equation

$$\mu_{min}(T(\lambda)) = 0, \tag{3.55}$$

where μ_{min} is the minimum eigenvalue of the positive semi-definite matrix $T(\lambda)$. Due to the truncation, however, the problem (3.54) generally does not have real-valued solutions for λ. Indeed, numerical tests suggest that T is typically strictly positive definite for all λ and then a non-zero solution for c does not exist. Figure 3.2 illustrates the dependence of the smallest eigenvalue $\mu_{min}(T(\lambda))$ of $T(\lambda)$ on λ.

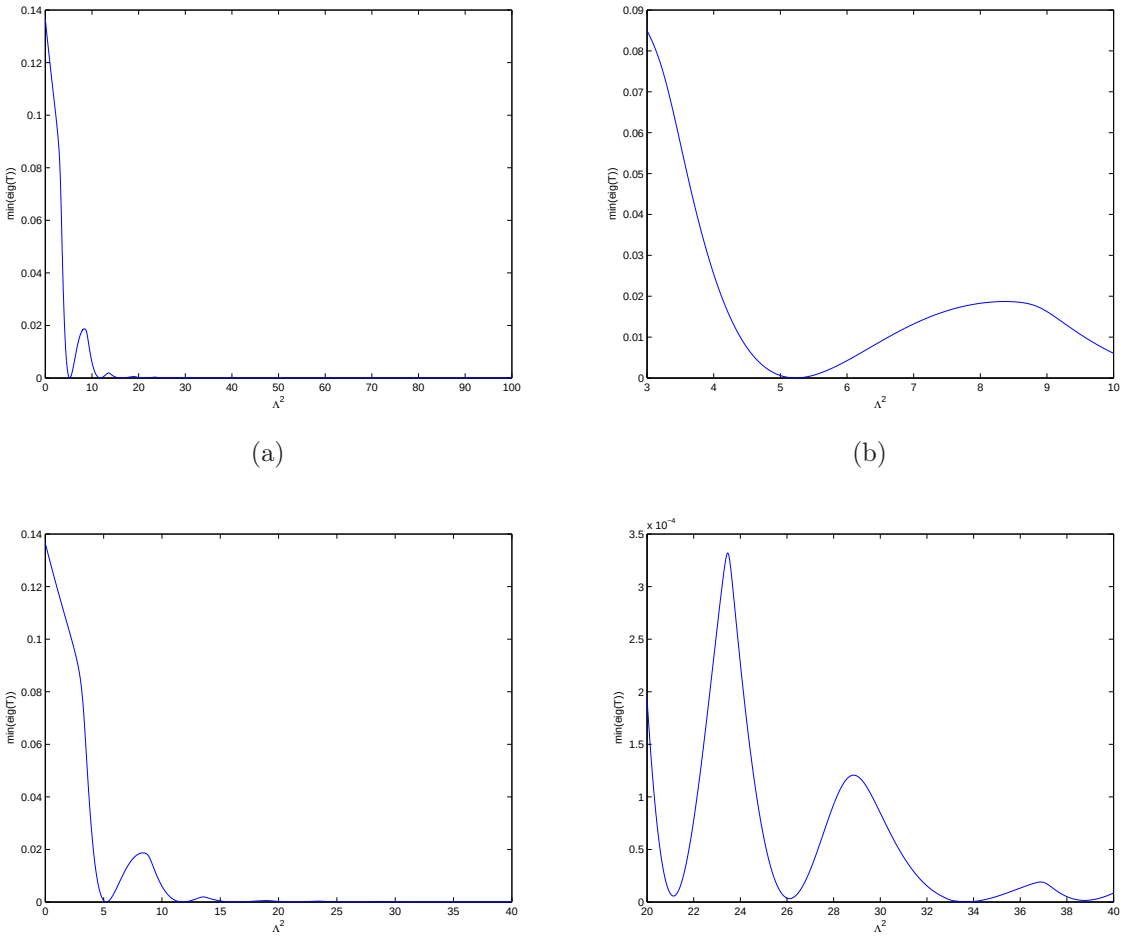


Figure 3.2: Minimum eigenvalue of matrix T: $\mu_{min}(T(\lambda))$ for $\rho_{xy} = 0.8, \rho_{xz} = 0.2, \rho_{yz} = 0.5$ and $N = 15$. Note the different plot ranges in (a) and (b).

In order to find an approximation to the original problem, we need to "regularize" it in the sense that there exists a solution of the truncated problem that is close to the solution of the infinite problem when the number of terms increases.

A first simple approach is based on the minima of μ_{min} (instead of its roots). Here, we use Chebyshev polynomials (Driscoll et al. (2014)) to approximate all local minima of the left-hand side of (3.55), and use these as approximation to the nonlinear eigenvalues.

More sophisticated and efficient methods by considering complex nonlinear eigenvalues, based on Chebyshev interpolation of $T(\lambda)$ directly and on contour integration, respectively, are given in Sections 3.6.6.2 and 3.6.6.3 below.[1]

3.6.6.2 Chebyshev polynomial interpolation of $T(\lambda)$

In this approach, we approximate $T(\lambda)$ for complex λ using Chebyshev polynomials,

$$T(\lambda) \approx \widehat{T}(\lambda) \equiv \sum_{j=0}^{m} A_j \tilde{T}_j(\lambda), \tag{3.56}$$

where $\{\tilde{T}_j, 0 \le j \le m\}$ is a (scalar) Chebyshev polynomial basis of degree m and $A_j \in \mathbb{R}^{N\times N}$ are the corresponding coefficients.

The following result is a modification of Theorem 18.1 in Trefethen (2013) (see, for example, Nakatsukasa et al. (2015), Section 7.2).

Theorem 3.1. *The solutions of the polynomial eigenvalue problem for $\widehat{T}$ in* (3.56) *are the eigenvalues of the linear matrix pencil (of dimension $mN \times mN$)*

$$\lambda \begin{bmatrix} A_m & & & & \\ & I_N & & & \\ & & \ddots & \ddots & \ddots \\ & & & & I_N \end{bmatrix} - \frac{1}{2} \begin{bmatrix} -A_{m-1} & I_N - A_{m-2} & -A_{m-3} & \cdots & A_0 \\ I_N & 0 & I_N & & \\ & \ddots & \ddots & \ddots & \\ & & I_N & 0 & I_N \\ & & & 2I_N & 0 \end{bmatrix}, \tag{3.57}$$

(i.e., all λ such that the above matrix is singular), where entries not displayed are zero.

Thus, by finding the eigenvalues of the linear matrix pencil we obtain the solution of a polynomial eigenvalue problem, which is an approximation to the nonlinear eigenvalue problem (3.54).

3.6.6.3 Contour integral method

The following method due to Beyn (2012) finds nonlinear complex eigenvalues within a given contour Γ. The main idea is to use the fact that

$$\frac{1}{2\pi i} \int_\Gamma f(z) T(z)^{-1} \, \mathrm{d}z = \sum_{n=1}^{n(\Gamma)} f(\lambda_n) v_n w_n^*,$$

[1] We avoid using $\int_0^{\bar{\omega}} f(\varphi)\bar{g}(\varphi) \, \mathrm{d}\varphi$ instead of $\langle \cdot, \cdot \rangle$ in (3.49) due to the positive definiteness.

where λ_j is the nonlinear eigenvalue, v_n and w_n^* are corresponding right and left eigenvectors.

Applying the last equation for $f(z) = 1$ and $f(z) = z$, we get

$$A_0 = \sum_n v_n w_n^* = VW^*,$$
$$A_1 = \sum_n \lambda_n v_n w_n^* = V\Lambda W^*.$$

From the last two equations, we can find eigenvalues Λ. We give the main steps in Algorithm 2 and subsequently discuss the choice of contour.

Algorithm 2 Contour integral algorithm for nonlinear eigenvalue problem

1: $l = 1$
2: Compute $A_0 = \frac{1}{2\pi i}\int_\Gamma T(z)^{-1}\hat{V}\,\mathrm{d}z$ and $A_1 = \frac{1}{2\pi i}\int_\Gamma zT(z)^{-1}\hat{V}\,\mathrm{d}z$ numerically.
3: Compute SVD $A_0 = V\Sigma W^H$.
4: Perform a rank test for Σ, i.e., find $0 < k \leq l$ such that

$$\sigma_1 \geq \sigma_2 \geq \ldots tol_{rank} > \sigma_{k+1} \approx \ldots \approx \sigma_l \approx 0.$$

5: **if** $k = l$ **then** $l = l + 1$ and go to line 1;
6: **else** let $V_0 = V(1:N, 1:k), W_0 = W(1:l, 1:k)$ and $\Sigma_0 = \mathrm{diag}(\sigma_1, \ldots, \sigma_k)$.
7: **end if**
8: Compute $B = V_0^H A_1 W_0 \Sigma_0^{-1}$.
9: Solve the eigenvalue problem for B: $BS = S\Lambda$, where $S = (v_1|v_2|\ldots|v_k)$.
10: For all j with $||T(\lambda_j)v_j|| \leq tol_{res}$ and $\lambda_j \in \mathrm{int}(\Gamma)$ accept λ_j as nonlinear eigenvalue and $v_j = V_0 s_j$ as eigenvector.

A few further specifications and comments:

1. In line 2, we choose $\hat{V} = (I_l|O)^T$ with I_l the $l \times l$ identity matrix and O an $l \times (N-l)$ matrix of zeros.

2. The integrands in line 3 are found by solution of linear systems $T(z)x = \hat{V}$ for x.

3. For the linear systems, SVDs and linear eigenvalue problems we use the Matlab default functions.

4. We choose as contours circles with the center μ and radius R. Then, for $\varphi(t) = \mu + R\exp\left(\frac{2\pi ij}{L}\right)$, A_0 and A_1 can be approximated by quadrature with L nodes,

$$A_{0,L} = \frac{R}{L}\sum_{j=0}^{L-1} T(\varphi(t_j))^{-1}\hat{V}\exp\left(\frac{2\pi ij}{L}\right),$$
$$A_{1,L} = \mu A_{0,L} + \frac{R^2}{L}\sum_{j=0}^{L-1} T(\varphi(t_j))^{-1}\hat{V}\exp\left(\frac{4\pi ij}{L}\right).$$

5. As μ can be large, for computational stability it is better to shift the coordinates to $\tilde{z} = z - \mu$.

3.6.7 Numerical tests

In this section, we discuss the application of the methods in Sections 3.6.2 and 3.6.6 for solving (3.54) and give numerical tests.

3.6.7.1 Matrix assembly

To compute the elements of T, we have to evaluate the integral in (3.49). We notice that

$$f_{mn}(\varphi') = \sin(k_m\varphi')\sin(k_n\varphi')\mathrm{e}^{(k_m+k_n)Z(\varphi')}{}_2F_1\left(\lambda, 1-\lambda, 1+k_m, \frac{\mathrm{e}^{2Z(\varphi')}}{1+\mathrm{e}^{2Z(\varphi')}}\right) \times \\ \times {}_2F_1\left(\lambda, 1-\lambda, 1+k_n, \frac{\mathrm{e}^{2Z(\varphi')}}{1+\mathrm{e}^{2Z(\varphi')}}\right)$$

is analytic on $\{|z|<1\}$ and $f'_{mn}(0) = f'_{mn}(\bar{\omega})$, where $\bar{\omega} = \arccos(-\rho_{xy})$. Then, according to Javed and Trefethen (2014), the trapezoidal quadrature rule with K nodes converges with order $O(K^{-4})$ with small constant.

An advantage of using the standard trapezoidal rule compared to other, adaptive quadrature rules is that on a fixed grid we can precompute the functions

$$f_n(\varphi') = \sin(k_n\varphi')\mathrm{e}^{k_nZ(\varphi')}{}_2F_1\left(\lambda, 1-\lambda, 1+k_n, \frac{\mathrm{e}^{2Z(\varphi')}}{1+\mathrm{e}^{2Z(\varphi')}}\right).$$

Therefore, we use only N instead of N^2 computations of hypergeometric functions, which is an expensive operation.

The elements of $T(\lambda)$ grow fast in n and m due to the term $\exp((k_n+k_m)Z(\varphi))$, and for large N can become highly inaccurate. We will use the following straightforward result to improve the condition of T.

Proposition 3.1. *Consider non-singular square matrices V and W. Define $\tilde{T}(\lambda) = VT(\lambda)W$. Then, λ is a nonlinear eigenvalue of $T(\lambda)$ if and only if it is a nonlinear eigenvalue of $\tilde{T}(\lambda)$.*

Define the matrices as $V = W = D$ with

$$D = \mathrm{diag}(\mathrm{e}^{-\alpha k_1}, \mathrm{e}^{-\alpha k_2}, \ldots, \mathrm{e}^{-\alpha k_N}).$$

Then, the elements of $\tilde{T}(\lambda) = DT(\lambda)D$ are

$$\tilde{T}(\lambda)_{mn} = \int_0^{\bar{\omega}} \sin(k_m\varphi')\sin(k_n\varphi')\mathrm{e}^{(k_m+k_n)(Z(\varphi')-\alpha)}{}_2F_1\left(\lambda, 1-\lambda, 1+k_m, \frac{\mathrm{e}^{2Z(\varphi')}}{1+\mathrm{e}^{2Z(\varphi')}}\right) \\ \times {}_2F_1\left(\lambda, 1-\lambda, 1+k_n, \frac{\mathrm{e}^{2Z(\varphi')}}{1+\mathrm{e}^{2Z(\varphi')}}\right) \mathrm{d}\varphi'. \quad (3.58)$$

Transformation (3.58) helps to decrease the elements proportionally to the growth rate, where α is chosen in the tests as maximum or average value of $Z(\varphi)$.

The standard Matlab hypergeometric function is very slow. Therefore, we implemented our own function in C and inserted it into Matlab as mex function.

3.6.7.2 Computation of eigenvalues

Now we consider the application of the numerical methods from Section 3.6.6 to the truncated problem (3.54).

Chebyshev approximation of smallest eigenvalue of $T(\lambda)$ (Section 3.6.6.1). For this method the application is straightforward, using Cholesky factorisation and *chebfun* (Driscoll et al. (2014)) to find the local minima.

Chebyshev polynomial interpolation (Section 3.6.6.2). Here, we consider complex eigenvalues and Chebyshev interpolation of $T(\lambda)$ is implemented using *chebfun* (Driscoll et al. (2014)). The resulting generalised linear eigenvalue problem is solved by the default method in Matlab.

Contour integral method (Section 3.6.6.3). To find all (complex) eigenvalues, we have to pick suitable centres and radii to apply Algorithm 2. Based on the empirical observation that the distribution of the eigenvalues is roughly uniform (see Table 3.1 below), we can choose a fixed radius and increase the centre by slightly under twice that amount to find the next eigenvalue. We discard any eigenvalues found previously due to the overlap of contours. This is summarised in Algorithm 3.

Algorithm 3 Contour integral algorithm for nonlinear eigenvalue problem

1: Choose $\mu = \mu_0$ and R as discussed above.
2: **while** $\mu \leq \mu_{max}$ **do**
3: Use Algorithm 2 with $\Gamma = C(\mu, R)$, a circle with centre μ and radius R.
4: $\mu = \mu + 2R$
5: **end while**

In Tables 3.1 and 3.2, we compare the approximate nonlinear eigenvalues and the computational time for different methods (from Section 3.6.6) in the semi-analytical approach with those from the finite element method in Lipton and Savescu (2014). Consider parameters $\rho_{xy} = 0.8, \rho_{xz} = 0.2, \rho_{yz} = 0.5$. For the semi-analytical method we take $N = 20$ for the approximation of the infinite sum in (3.50), $K = 1000$ and $L = 500$ quadrature nodes.

EV	Finite elements	Chebyshev approximation of smallest eigenvalue of $T(\lambda)$	Contour integrals / Chebyshev interpolation
Λ_1^2	5.232	5.228	5.229 (0.028)
Λ_2^2	11.805	11.787	11.787 (0.076)
Λ_3^2	16.313	16.285	16.284 (0.068)
Λ_4^2	21.204	21.149	21.147 (0.151)
Λ_5^2	26.184	26.112	26.109 (0.188)
Λ_6^2	33.392	33.230	33.261 (0.215)
Λ_{15}^2	72.911	72.841	72.708 (0.731)
Λ_{30}^2	138.087	139.421	139.391 (1.331)

Table 3.1: Eigenvalues for parameters $\rho_{xy} = 0.8, \rho_{xz} = 0.2, \rho_{yz} = 0.5$. In the last column, the imaginary part is in brackets. As the contour integral method and Chebyshev polynomial interpolation gave identical results up to 5 digits, we unite them in one column.

Chebyshev approximation of smallest eigenvalue of $T(\lambda)$	Chebyshev interpolation of $T(\lambda)$	Contour integrals
561.17 s	60.27 s	220.79 s

Table 3.2: Computation time of first 30 eigenvalues.

We also compare the numerically computed nonlinear eigenvalues to the analytic expressions in the case $\rho_{xy} = \rho_{xz} = \rho_{yz} = 0$ (see Appendix 3.6.4.1), in Table 3.3. We can see that the semi-analytical method is vastly superior to the finite element method. The semi-analytical method has the advantage in this case that there exists a finite basis and the truncation for the nonlinear eigenvalue problem after sufficiently many, say N terms does not change the solution (see Section 3.6.4.1). Here, N is chosen large enough so that the remaining errors come predominantly from the approximate solution of the linear eigenvalue problem in the case of the Chebyshev method and the singular value decomposition in the contour integral method.

EV	Value	Finite elements	Chebyshev approx. of sm. eig of $T(\lambda)$	Chebyshev interpolation	Contour integrals
Λ_1^2	12.0	0.018	4.1×10^{-15}	3.6×10^{-12}	1.3×10^{-12}
Λ_2^2	30.0	0.097	9.5×10^{-14}	1.1×10^{-11}	3.4×10^{-12}
Λ_3^2	30.0	0.105	9.5×10^{-14}	1.1×10^{-11}	3.4×10^{-12}
Λ_4^2	56.0	0.309	7.7×10^{-12}	6.0×10^{-11}	6.1×10^{-11}
Λ_5^2	56.0	0.353	7.7×10^{-12}	6.0×10^{-11}	6.1×10^{-11}
Λ_7^2	90.0	0.767	1.0×10^{-10}	5.4×10^{-10}	7.8×10^{-10}
Λ_{15}^2	132.0	2.311	2.3×10^{-5}	1.8×10^{-9}	4.5×10^{-8}
Λ_{30}^2	306.0	8.896	1.3×10^{-2}	4.1×10^{-8}	1.0×10^{-6}

Table 3.3: Error of eigenvalues for $\rho_{xy} = 0.0, \rho_{xz} = 0.0, \rho_{yz} = 0.0$. We present the difference between eigenvalues computed by different methods and the analytical solution.

From the numerical experiments we can conclude that the method based on the smallest linear eigenvalue is slow. The methods based on contour integrals and Chebyshev interpolation show the best results. From a computational point of view, Chebyshev interpolation is faster than the contour integral method, because for the latter we have to run the computations several times on different circles, as well as compute numerical integrals along contours.

The complexity of the Chebyshev interpolation method consists of two steps. The first step is the interpolation of $T(\lambda)$ in λ using Chebyshev polynomials as in (3.56) to obtain a polynomial eigenvalue problem. This can be performed in $O(N^2m)$, where N is the number of terms in the eigenvalue expansion and m is the dimension of the Chebyshev basis. The second step is to solve the polynomial eigenvalue problem, for which the complexity equals that of the linear eigenvalue decomposition of a matrix of size $Nm \times Nm$. Thus, the complexity of the second step is proportional to that of matrix multiplication, i.e., $O(N^3m^3)$ in practice.[2]

We give an empirical analysis of the CPU time[3] in Table 3.4. Since the first step has a higher constant factor, it dominates for small m, while for large m the second step dominates.

Once we have computed the eigenvalues and the corresponding eigenvectors, we can compute $\Psi(\varphi, \zeta)$ using expression (3.47). An example of eigenvectors is given in Figure 3.3, shown in the (φ, θ)-plane.

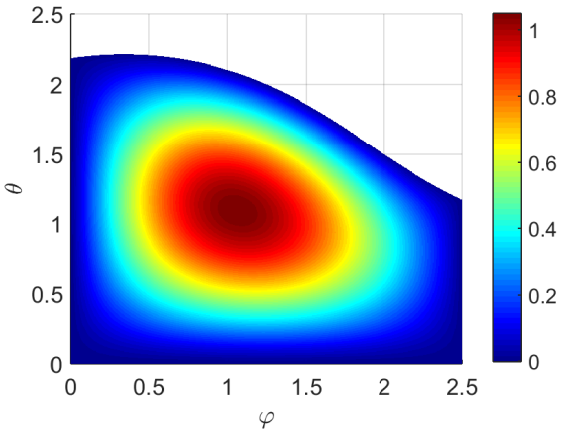
3.6.7.3 Convergence of expansion and computational speed

We now focus on the most efficient method, the Chebyshev interpolation method from Section 3.6.6.2. To explore the convergence numerically, we compare our results for the Green's function with the analytical solution available for the special correlation case $\rho_{xy} = -\cos\left(\frac{\pi}{3}\right), \rho_{xz} = -\cos\left(\frac{\pi}{3}\right), \rho_{yz} = 0$ using a method of images (Escobar et al. (2013)).

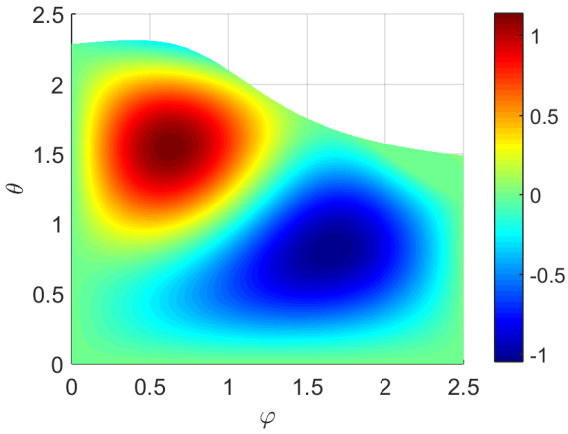
It is worth pointing out that we should not expect uniform convergence in (t', x', y', z'), especially as $t' \downarrow t$, because the initial condition is a Dirac delta. In applications, the quantity of interest is usually an integral in the space variables of the Green's function for $t' > 0$. It is well-known that for $t' > 0$ the Green's function is smoothed. In our experiments, we consider the spatial L_2 norm (for fixed t') for comparison.

[2]Theoretically, algorithms with $O(N^\omega m^\omega)$ for $2 < \omega < 2.37$ are available (see Demmel et al. (2007)), but the constant is found to be too large to be competitive for the matrix size relevant here.

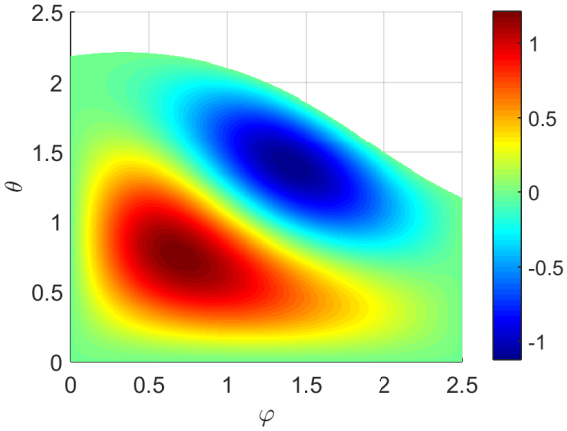
[3]We performed the calibration in Matlab on a PC with Intel Core i7 3.4Ghz (4 cores) CPU.



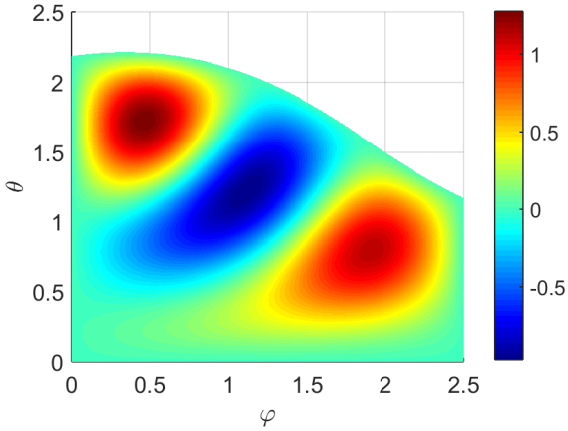
(a) Eigenvector 1: $\Lambda_1^2 = 5.2$



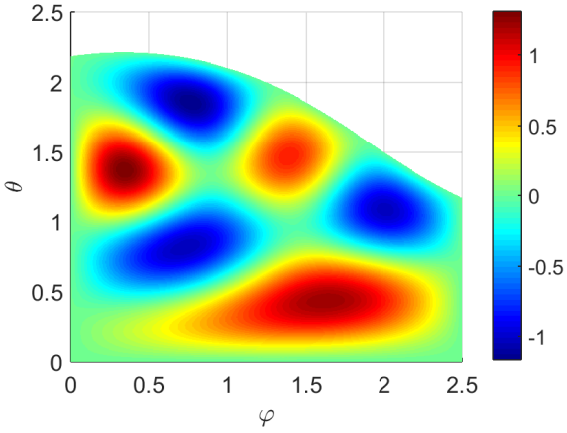
(b) Eigenvector 2: $\Lambda_2^2 = 11.8$



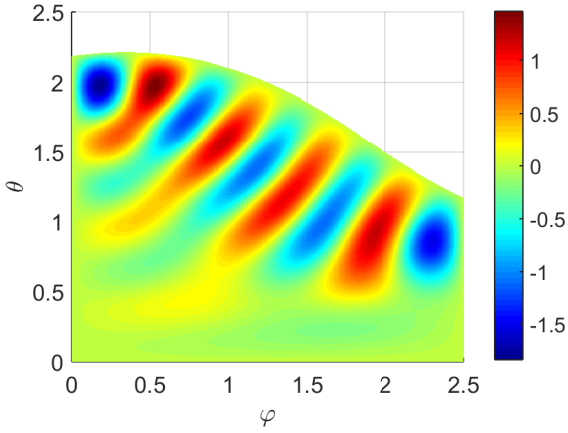
(c) Eigenvector 3: $\Lambda_3^2 = 16.3$



(d) Eigenvector 4: $\Lambda_4^2 = 21.1$



(e) Eigenvector 8: $\Lambda_8^2 = 39.3$



(f) Eigenvector 30: $\Lambda_{30}^2 = 139.4$

Figure 3.3: An example of eigenvectors and corresponding eigenvalues for the domain obtained for $\rho_{xy} = 0.8, \rho_{xz} = 0.2, \rho_{yz} = 0.5$.

Our truncated Green's function approximation has the form

$$G(\tau, r', \varphi', \theta', r, \varphi, \theta) = \frac{\mathrm{e}^{-\frac{r'^2+r^2}{2\tau}}}{\tau\sqrt{r'r}} \sum_{l=1}^{M} I_{\lambda_l}\left(\frac{r'r}{\tau}\right) \times$$
$$\times \left(\sum_{n=1}^{N} c_n^l \sin(k_n\varphi') \tan^{k_n}\left(\frac{\theta'}{2}\right) {}_2F_1\left(\lambda_l, 1-\lambda_l, 1+k_n, \sin^2\frac{\theta'}{2}\right)\right) \times$$
$$\times \left(\sum_{n=1}^{N} c_n^l \sin(k_n\varphi) \tan^{k_n}\left(\frac{\theta}{2}\right) {}_2F_1\left(\lambda_l, 1-\lambda_l, 1+k_n, \sin^2\frac{\theta}{2}\right)\right).$$

We have several numerical parameters:

(i) M, the number of terms in the superposition solution of (3.34), i.e., the number of nonlinear eigenvalues and eigenvectors in (3.50) we approximate;

(ii) N, the dimension of the nonlinear eigenvalue problem (3.54) approximating the infinite-dimensional problem (3.50);

(iii) m, the dimension of the Chebyshev basis in (3.56) to approximate the nonlinear eigenvalue problem by a polynomial one;

(iv) K, the number of quadrature points to approximate the integral in (3.49); see also Section 3.6.7.1.

We expect convergence of the approximation as those four parameters go to infinity. This involves in particular that c_l^n, combined with the terms they are multiplied by, go to zero sufficiently fast as $n, l \to 0$. Moreover, one needs that for fixed n, l the c_l^n and λ_l, which implicitly depend on N through the truncation of the infinite-dimensional eigenvalue problem and on m and K by the approximations made to the finite-dimensional eigenvalue problems, converge as $N, m, K \to \infty$.

In absence of theoretical error estimates, we test this empirically. In our experiments, we choose $K = 1000$ nodes for the trapezoidal rule for (3.49); very good tolerance is already achieved with much fewer nodes, however, even for $K = 1000$ the computational time for these integrals is not very significant (about 10-15% of the total eigenvalue computation procedure).

We now provide three experiments: in each, we fix two of the remaining parameters (M, N, m) to be large enough and vary the third parameter to compare the Green's function with the analytical solution. For $N = 30, M = 50, m = 100$, in our experiments, further increase of either parameter does not affect the Green's function's value beyond the accuracy required here.

The results and computational times are given in Table 3.4 below. Recall that the infinite-dimensional nonlinear eigenvalue problem (3.50) is ultimately approximated by the (mN)-dimensional generalized linear eigenvalue problem for (3.57). To investigate the convergence in the number of terms in the expansion, M, we choose the M smallest out of the resulting mN eigenvalues and corresponding eigenvectors. As the main computational

cost comes from the solution of the eigenvalue problem, which is here identical for all M, we do not report computational times for different M.

In comparison, using the finite element method from Lipton and Savescu (2014) with a 1500 point mesh, we get the precision 3.1×10^{-5} with CPU time 621s.[4] This accuracy is comparable with that of the semi-analytical solution for m around 15 and N around 7, which is computable within a fraction of a second. Note that the approach of Lipton and Savescu (2014) requires the solution of a 1500-dimensional (determined by the number of finite elements) generalized linear eigenvalue problem. In our tests, however, the cost was dominated by the construction of the locally refined finite element mesh.

m	10	20	50	80	100
Error	2.52×10^{-4}	7.73×10^{-6}	6.21×10^{-7}	1.76×10^{-7}	8.60×10^{-8}
CPU time (s)	79.9	94.8	256.5	604.1	1070.1
N	5	10	20	25	30
Error	7.92×10^{-5}	3.46×10^{-6}	2.68×10^{-7}	8.60×10^{-8}	8.60×10^{-8}
CPU time (s)	9.9	56.6	333.1	667.4	1070.1
M	5	10	20	40	50
Error	4.87×10^{-4}	1.19×10^{-5}	9.71×10^{-7}	1.02×10^{-7}	8.60×10^{-8}

Table 3.4: L_2 norm of the Green's function's error for different m ($N = 30, M = 50$), N ($m = 100, M = 50$), and M ($m = 100, N = 30$).

As a further benchmark, we present the results for a second order ADI finite-difference method (the Hundsdorfer-Verwer scheme; see Hundsdorfer and Verwer (2013)) in Table 3.5 below. We use up to $N_X = 100$ points in each spatial direction and $N_T = 200$ points in time.

The numerical initial condition is chosen as

$$G(0, x_1^i, x_2^j, x_3^k, x_1, x_2, x_3) = \begin{cases} \frac{1}{\Delta x_1 \Delta x_2 \Delta x_3}, & x_1^i = x_1, x_2^j = x_2, x_3^k = x_3, \\ 0, & \text{otherwise.} \end{cases}$$

in order to approximate the Dirac delta, with the mesh constructed in such a way that (x_1, x_2, x_3) coincides with a mesh point. This is a special case of the approximations to non-smooth initial conditions analysed numerically in Pooley et al. (2003) and by means of Fourier methods in Reisinger and Whitley (2014).

The data in Table 3.5 show second order convergence in both N_X and N_T.

[4]We performed the calibration in Matlab on a PC with Intel Core i7 3.4Ghz (4 cores) CPU.

N_X	25	50	100
Error	4.44×10^{-4}	1.04×10^{-4}	2.65×10^{-5}
CPU time (s)	30.8	274.3	2239.1
N_T	50	100	200
Error	3.41×10^{-4}	9.23×10^{-5}	2.65×10^{-5}
CPU time (s)	723.2	1156.6	2239.1

Table 3.5: L_2 norm of the Green's function's error for the ADI method with different N_X ($N_T = 200$) and N_T ($N_X = 100$).

This test demonstrates the superior accuracy of the semi-analytical solution for relatively few terms in comparison to the ones obtained for the ADI scheme on the highest computationally feasible levels.

It is worth pointing out though that the semi-analytical expression gives the density in a single point, while the finite difference solution is simultaneously obtained on a whole mesh. This is useful, for example, to compute survival probabilities, where one needs to integrate the density.

3.7 Extended Default Model with one bank

For $N = 1$, (2.58)–(2.60) become

$$\frac{\partial V}{\partial t} + \mathcal{L}^{(1)} V = \chi(t, x, y, z), \tag{3.59}$$

$$V(t, 0) = \phi_0(t), \quad V(t, x) \underset{x \to +\infty}{\longrightarrow} \phi_\infty(t), \tag{3.60}$$

$$V(T, x) = \psi(x), \tag{3.61}$$

with suitably defined terminal conditions ψ, boundary conditions ϕ and right-hand sides χ, and where the Kolmogorov backward operator is

$$\mathcal{L}^{(1)} f = \frac{1}{2} f_{xx} + \xi f_x.$$

3.7.1 Survival probability

Using Green's formula (3.3) and taking into account the drift, the survival probability (A.11) is given by

$$Q(t, x) = \int_{\mu^=}^{+\infty} G(T - t, x', x) \, \mathrm{d}x'. \tag{3.62}$$

Hence, applying (3.14), we get

$$Q(t, x) = N\left(-\frac{\mu^= - x - \xi(T - t)}{\sqrt{T - t}}\right) - \mathrm{e}^{-2\xi x} N\left(-\frac{\mu^= - x + \xi(T - t)}{\sqrt{T - t}}\right), \tag{3.63}$$

where $N(x)$ is normal CDF.

3.7.2 Credit Default Swap

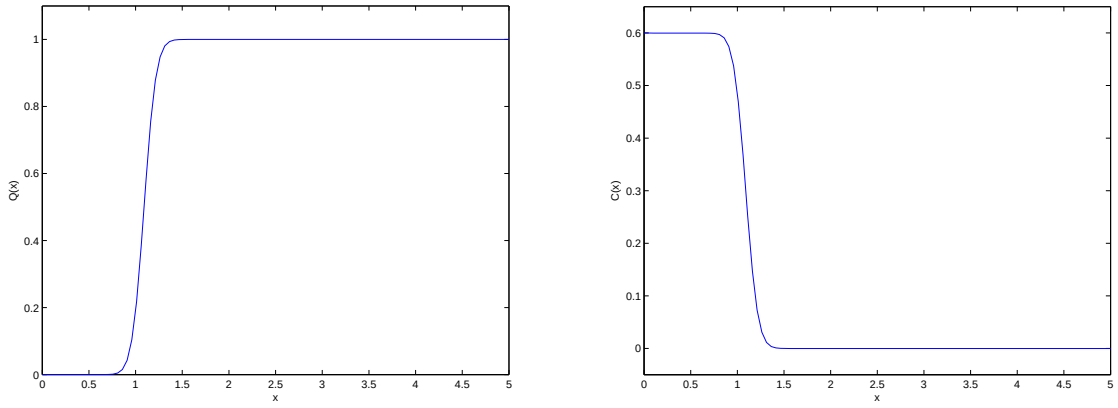
Using Green's formula (3.3) and taking into account the drift for (A.13), we get

$$\begin{aligned} C(t,x) = (1-R)\int_0^{\mu^=} G(T-t,x',x)\,\mathrm{d}x' + \frac{1-R}{2}\int_t^T \frac{\partial G(T-\tau,0,x)}{\partial x'}\,\mathrm{d}\tau \\ - c\int_t^T \int_0^{+\infty} G(T-\tau,x',x)\,\mathrm{d}\tau\mathrm{d}x'. \end{aligned} \tag{3.64}$$

The integrals in the last equation can be computed analytically and expressed via the normal CDF.

3.7.3 Numerical examples

Consider $\sigma = 1, R = 0.4, L = 70, T = 1$. In Figure 3.4 we present survival probability and CDS spreads computed using the methods described above.



(a) (b)

Figure 3.4: The model with one bank: (a) Survival probability (b) CDS spread.

3.8 Extended Default Model with three banks

For $N = 3$, (2.58)–(2.60) become

$$\frac{\partial V}{\partial t} + \mathcal{L}^{(3)}V = \chi(t,x,y,z), \tag{3.65}$$

$$V(t,0,y,z) = \phi_{0,x}(t,y,z), \quad V(t,x,y,z) \underset{x\to+\infty}{\longrightarrow} \phi_{\infty,x}(t,y,z), \tag{3.66}$$

$$V(t,x,0,z) = \phi_{0,y}(t,x,z), \quad V(t,x,y,z) \underset{y\to+\infty}{\longrightarrow} \phi_{\infty,y}(t,x,z), \tag{3.67}$$

$$V(t,x,y,0) = \phi_{0,z}(t,x,y), \quad V(t,x,y,z) \underset{z\to+\infty}{\longrightarrow} \phi_{\infty,z}(t,x,y), \tag{3.68}$$

$$V(T,x,y,z) = \psi(x,y,z), \tag{3.69}$$

with suitably defined terminal conditions ψ, boundary conditions ϕ and right-hand sides χ, and where the Kolmogorov backward operator is

$$\mathcal{L}^{(3)}f = \frac{1}{2}f_{xx} + \frac{1}{2}f_{yy} + \frac{1}{2}f_{zz} + \rho_{12}f_{xy} + \rho_{13}f_{xz} + \rho_{23}f_{yz} + \xi_1 f_x + \xi_2 f_y + \xi_3 f_z.$$

3.8.1 Joint survival probability

The corresponding Kolmogorov backward equation for the joint survival probability is

$$\begin{aligned}
&\frac{\partial Q}{\partial t} + \mathcal{L}^{(3)} Q = 0, \\
&Q(t,0,y,z) = 0, \quad Q(t,x,0,z) = 0, \quad Q(t,x,y,z) = 0, \\
&Q(T,x,y,z) = \mathbb{1}_{(x,y,z)\in D_{xyz}},
\end{aligned}$$

where $D_{xyz} = \{x \geq \mu_x, y \geq \mu_y, z \geq \mu_z\}$ is the subset of $\mathbb{R}^3_+$ where all banks survive.

Due to various regulation requirements, assets of large banks cannot drop below their liabilities, which means that their recovery rates R should be close to 1. This implies that the domain D_{xyz} becomes the positive octant $\mathbb{R}^3_+$.

The analytical solution for the survival probability in the positive quadrant (i.e., two dimensions) with non-zero drift was recently found by Itkin and Lipton (2017). Here, we extend this result to three dimensions using the expression for the Green's function from Section 3.6.2 and an expansion in Gegenbauer polynomials.

In this section alone, we use (x,y,z) instead of (x',y',z'), and (x_0,y_0,z) instead of (x,y,z), to shorten the notation in the integrals, and similarly for (r,φ,θ). The solution Q can then be written as

$$Q(\tau,x_0,y_0,z_0) = \iiint\limits_{\mathbb{R}^3_+} G(\tau,x',y',z',x_0,y_0,z_0)\,\mathrm{d}x'\mathrm{d}y'\mathrm{d}z'.$$

Using (3.53),

$$\begin{aligned}
Q(\tau,r_0,\varphi_0,\theta_0) = &\int_0^{+\infty}\int_0^{\bar{\omega}}\int_0^{\Theta(\varphi)} \exp\left(\xi_\alpha(r\sin\theta\sin\varphi - r_0\sin\theta_0\sin\varphi_0) + \right. \\
&\left. \xi_\beta(r\sin\theta\cos\varphi - r_0\sin\theta_0\cos\varphi_0) + \xi_\gamma(r\cos\theta - r_0\cos\theta_0)\right) G(\tau,r,\theta,\varphi) r^2\sin\theta\,\mathrm{d}\theta\mathrm{d}\varphi\mathrm{d}r = \\
&\frac{\mathrm{e}^\kappa}{\tau\sqrt{r_0}} \sum_{l=1}^{\infty} \Psi_l(\varphi_0,\theta_0) \int_0^{\bar{\omega}}\int_0^{\Theta(\varphi)} \Psi_l(\varphi,\theta)\sin\theta \int_0^{+\infty} \mathrm{e}^{-\alpha r^2} r^{\frac{3}{2}} \mathrm{e}^{\gamma(\varphi,\theta)r} I_{\lambda_l}(\beta r)\,\mathrm{d}r\mathrm{d}\theta\mathrm{d}\varphi, \quad (3.70)
\end{aligned}$$

where $\lambda_l = \sqrt{\Lambda_l^2 + \frac{1}{4}}$, $\alpha = \frac{1}{2\tau}$, $\beta = \frac{r_0}{\tau}$, $\gamma(\varphi,\vartheta) = \xi_\alpha\sin\theta\sin\varphi + \xi_\beta\sin\theta\cos\varphi + \xi_\gamma\cos\theta$ and

$$\kappa = -\xi_\alpha r_0\sin\theta_0\sin\varphi_0 - \xi_\beta r_0\sin\theta_0\cos\varphi_0 - \xi_\gamma r_0\cos\theta_0 - \alpha r_0^2.$$

Next, we use the idea of exponent representation via ultra-spherical polynomials from Itkin and Lipton (2017),

$$\mathrm{e}^{-Sx} = \Gamma(\nu)\left(\frac{S}{2}\right)^{-\nu} \sum_{k=0}^{\infty} (-1)^k(\nu+k) I_{\nu+k}(S) C_k^\nu(x),$$

where $C_k^\nu(x)$ are Gegenbauer polynomials and the parameter ν can be chosen arbitrarily.

Substituting this representation with $S = \beta R, x = -\gamma(\varphi,\theta)/\beta$ into (3.70), we get

$$Q(\tau, r_0, \varphi_0, \theta_0) = \frac{\mathrm{e}^{\kappa}}{\tau\sqrt{r_0}}\Gamma(\nu)\left(\frac{\beta}{2}\right)^{-\nu}\sum_{l=1}^{\infty}\sum_{\mu=0}^{\infty}(-1)^{\mu}(\mu+\nu)\Psi_l(\varphi_0,\theta_0)\times$$
$$\times\iint\limits_{\Omega}\Psi_l(\varphi,\theta)\sin\theta C_{\mu}^{\nu}(-\gamma(\varphi,\theta)/\beta)\,\mathrm{d}\theta\mathrm{d}\varphi\int_0^{+\infty}\mathrm{e}^{-\alpha r^2}r^{\frac{3}{2}-\nu}I_{\lambda_l}(\beta r)I_{\mu+\nu}(\beta r)\,\mathrm{d}r. \quad (3.71)$$

For simplicity we choose $\nu = 1$ and then use the identity (Gradštejn et al. (2007)) for the second integral

$$I_2^{l,\mu} = \int_0^{\infty}\mathrm{e}^{-\alpha r^2}r^{\frac{1}{2}}I_{\lambda_l}(\beta r)I_{\nu+\mu}(\beta r)\mathrm{d}r = 2^{-\lambda_l-\mu-2}\alpha^{-(\lambda_l+\mu+5/2)/2}\beta^{\lambda_l+\mu+1}$$
$$\cdot\frac{\Gamma[(3/2+\mu+\lambda_l)/2]}{\Gamma(\mu+1)\Gamma(\lambda_l+1)}{}_3F_3\left[\begin{array}{ccc}\frac{\lambda_l+\mu+2}{2} & \frac{\lambda_l+\mu+3}{2} & \frac{\lambda_l+\mu+5/2}{2}\\ \mu+2 & \lambda_l+1 & \mu+\lambda_l+2\end{array};\frac{\beta^2}{\alpha}\right], \quad (3.72)$$

where ${}_3F_3(a_1,b_1,c_1;a_2,b_2,c_2;z)$ is a generalized hypergeometric function.

Moreover, with $\nu = 1$ the Gegenbauer polynomials become the Chebyshev polynomials of the second kind, which admit the representation (Abramowitz and Stegun (1964))

$$U_l(x) = \sum_{k=0}^{[l/2]}(-1)^k C_k^{l-k}(2x)^{l-2k},$$

where $[x]$ is the floor function. Thus, the integral in (3.71) over Ω can be represented as

$$I_1^{l,\mu} = \sum_{k=0}^{[\mu/2]}2^{\mu-2k}(-1)^{\mu-k}C_k^{\mu-k}\iint\limits_{\Omega}\Psi_l(\varphi,\theta)\sin\theta(\xi_\alpha\sin\theta\sin\varphi+\xi_\beta\sin\theta\cos\varphi+\xi_\gamma\cos\theta)^{\mu-2k}\mathrm{d}\theta\mathrm{d}\varphi. \quad (3.73)$$

As a result, we get the expression for the integral

$$Q(\tau, r_0, \varphi_0, \theta_0) = \frac{2\mathrm{e}^{\kappa}}{\beta\tau\sqrt{r_0}}\sum_{l=1}^{\infty}\sum_{\mu=0}^{\infty}(-1)^{\mu}(\mu+1)\Psi_l(\varphi_0,\theta_0)\cdot I_1^{l,\mu}\cdot I_2^{l,\mu}, \quad (3.74)$$

where the expression for $I_1^{l,\mu}$ is given in (3.73) and the expression for $I_2^{l,\mu}$ in (3.72).

From a computational point of view, this semi-analytical representation has an advantage in comparison to a direct integration of the Green's function. Consider parameters M and μ_{max}, the limits in the summation (3.74), and N the number of terms in the eigenvalue expansion of Ψ_l. The integrals over Ω can be computed numerically. Since $\Theta(\theta)$ is smooth, we can transform the domain to a rectangle by a smooth change of variables and use a high order integration method such as Gaussian quadrature. We denote by N_Ω the number of two-dimensional integration nodes. First, we precompute Ψ_l for all l on the grid; it can be done with $O(MNN_\Omega)$ operations. Then, we precompute the integrals in (3.73), the complexity of which is is $O(\mu_{max}MN_\Omega)$. Then, we precompute $I_1^{l,\mu}$, with complexity $O(\mu_{max}^2 M)$. Finally, we compute (3.74), using our precomputations, with $O(\mu_{max}M)$ operations. As a result, overall complexity is $O(MNN_\Omega + \mu_{max}MN_\Omega + M\mu_{max}^2)$. In comparison, the direct integration of the Green's function with pre-computations of Ψ_l for all

l gives $O(MNN_\Omega + MN_rN_\Omega)$ operations, where N_r is the number of quadrature nodes for the variable r, and which requires a truncation at some r_{max} in order to deal with the improper integral over r. Note that (3.53) is for the de-drifted Green's function, so that after multiplication by the factor in (3.7) each term in the sum is a non-trivial function of (r', ϕ', θ') and no separation structure can be exploited. Usually, the values of μ_{max} and N can be chosen relatively small in comparison to N_r for comparable accuracy.

In our computations we use the parameters in Table 3.6, and liabilities as in Table 3.7.

σ_1	σ_2	σ_3	ρ_{xy}	ρ_{xz}	ρ_{yz}	R_1	R_2	R_3	T
1	1	1	0.8	0.2	0.5	0.4	0.45	0.4	1

Table 3.6: Assumed model parameters.

L_1	L_2	L_3	L_{12}	L_{13}	L_{21}	L_{23}	L_{31}	L_{32}
60	70	65	10	15	10	10	5	10

Table 3.7: Assumed external and mutual liabilities.

In Figure 3.5 we show the computed joint survival probabilities as a projection onto the y-z plane for different values of x.

3.8.2 Computation of Credit and Debt Value Adjustments

In this section, we compute valuation adjustments for credit default swaps (CDS). We associate x

Rewriting (A.43) in (x, y, z) variables, our PDE for CVA becomes

$$\begin{aligned}
&\frac{\partial V^{CVA}}{\partial t} + \mathcal{L}V^{CVA} = 0,\\
&V^{CVA}(t, 0, y, z) = (1 - R_{PS})V^{CDS}(t, z)_+,\\
&V^{CVA}(t, x, 0, z) = 0, \quad V^{CVA}(t, x, y, 0) = 0,\\
&V^{CVA}(T, x, y, z) = 0.
\end{aligned}$$

Then, by an application of Green's formula as in Lipton and Savescu (2014),

$$V^{CVA}(0, r, \varphi, \theta) = \frac{1 - R_{PS}}{2} \int_0^T \int_0^\infty \int_0^{\Theta(0)} \frac{V^{CDS}(t, r', 0, \theta')_+ G_{\varphi'}(t, r', 0, \theta', r, \varphi, \theta)}{\sin\theta'} \, \mathrm{d}\theta' \mathrm{d}r' \mathrm{d}t, \tag{3.75}$$

where V^{CDS} can be found analytically (Itkin and Lipton (2017)).

Similarly, the PDE for DVA in (x, y, z) variables becomes

$$
\begin{aligned}
&\frac{\partial V^{DVA}}{\partial t} + \mathcal{L}V^{DVA} = 0,\\
&V^{DVA}(t,0,y,z) = (1-R_{PS})V^{CDS}(t,z)_-,\\
&V^{DVA}(t,x,0,z) = 0, \quad V^{DVA}(t,x,y,0) = 0,\\
&V^{DVA}(T,x,y,z) = 0.
\end{aligned}
$$

Thus,

$$
V^{DVA}(0,r,\varphi,\theta) = \frac{R_{PB}-1}{2}\int_0^T\int_0^\infty\int_0^{\Theta(0)} \frac{V^{CDS}(t,r',0,\theta')_- G_{\varphi'}(t,r',\bar{\omega},\theta',r,\varphi,\theta)}{\sin\theta'}\,\mathrm{d}\theta'\mathrm{d}r'\mathrm{d}t. \tag{3.76}
$$

A distinct advantage of our semi-analytical method is that we can compute the derivative G_φ in (3.75) and (3.76) analytically as

$$
\begin{aligned}
G_{\varphi'}(\tau,r',0,\theta',r,\varphi,\theta) = \kappa(r',0,\theta',r,\varphi,\theta)\frac{\mathrm{e}^{-\frac{r'^2+r^2}{2\tau}}}{\tau\sqrt{r'r}}\sum_{l=1}^{\infty} I_{\lambda_l}\left(\frac{r'r}{\tau}\right)\times\\
\times\left(\sum_{n=1}^{\infty} c_n^l k_n \tan^{k_n}\left(\frac{\theta'}{2}\right) {}_2F_1\left(\lambda_l, 1-\lambda_l, 1+k_n, \sin^2\frac{\theta'}{2}\right)\right)\times\\
\times\left(\sum_{n=1}^{\infty} c_n^l \sin\left(k_n\varphi\right)\tan^{k_n}\left(\frac{\theta}{2}\right) {}_2F_1\left(\lambda_l, 1-\lambda_l, 1+k_n, \sin^2\frac{\theta}{2}\right)\right).
\end{aligned}
$$

Therefore, all terms in (3.75) and (3.76) can be computed analytically, and CVA can be computed by integration. As the functions are smooth, this can be done with a higher order integration method, such as Gaussian quadrature.

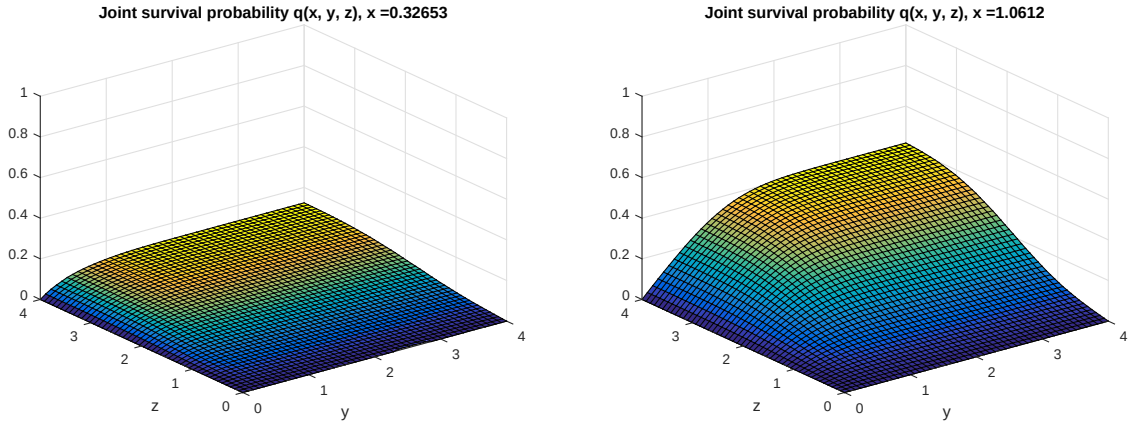
In our computations we use the parameters from Tables 3.6 and 3.7. In Figure 3.6, we then show the computed CVA as a projection on the z-x plane for different values of y.

The CVA decreases with the distance x (or x_{PS}) from the default boundary of the protection seller: the less likely the protection seller is to default, the lower the CVA. In terms of the distance to default z (or x_{RN}) of the reference name, the CVA is highest for intermediate values: if the RN is close to default, it is unlikely that the PS defaults before their CDS liability and the CVA is small; if the RN is far from default, it is unlikely that the protection is payable and again the CVA is small. The CVA mostly increases with the value of the protection buyer y (or x_{PB}) as their default, which is more likely for small y, voids the protection payment.

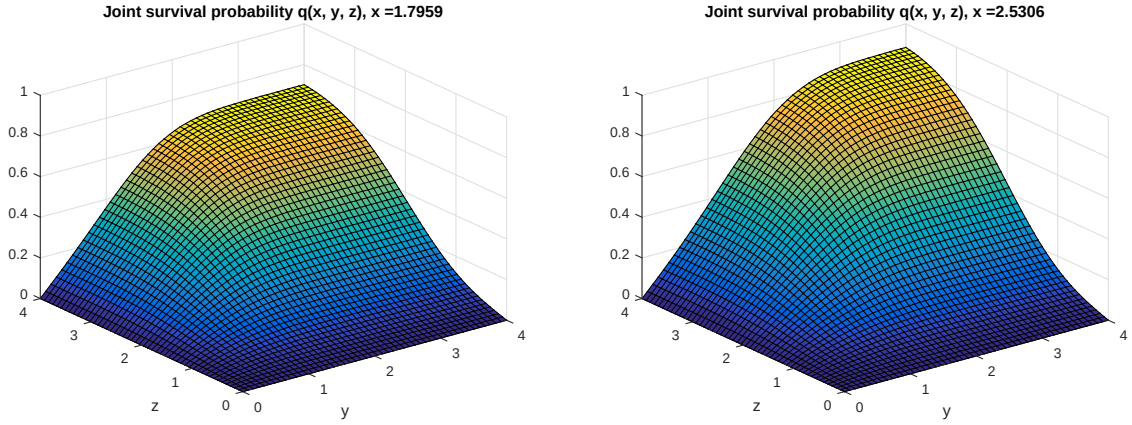
For all three firms, the marginal default probability decreases with the (rescaled, logarithmic) firm value. It increases with the firm's volatility. The dependence of the CVA on the marginal default probabilities and volatilities is therefore implicitly given by the plots in Figure 3.6 and we do not show them explicitly for brevity.

The impact of CVA and DVA on CDS prices in this model and its sensitivity to input parameters, specifically and importantly the correlations between the firms, is

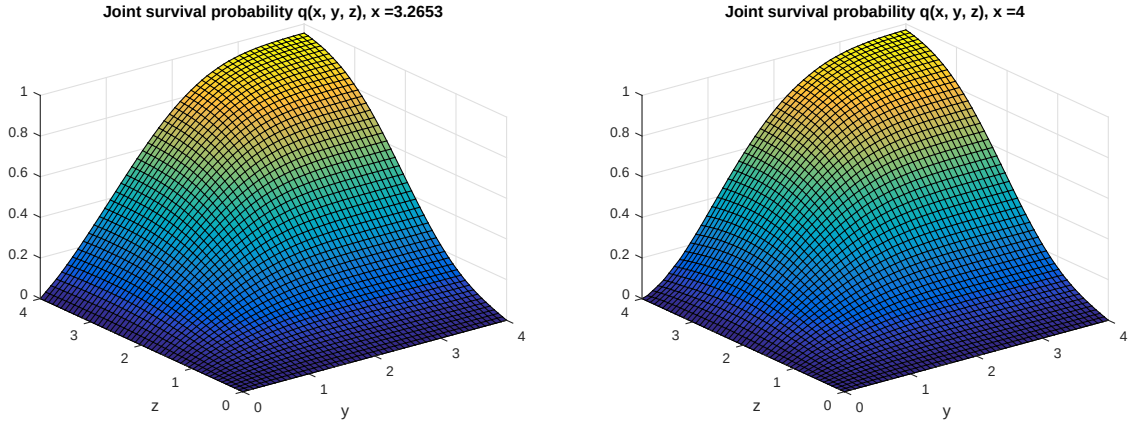
discussed in more detail in Section 7 of Lipton and Savescu (2014) (using the finite element approach). The complexity of the integration remains the same in the finite-element method, but, using our semi-analytical method, we compute derivatives in (3.75) and (3.76) analytically.



(a) (b)

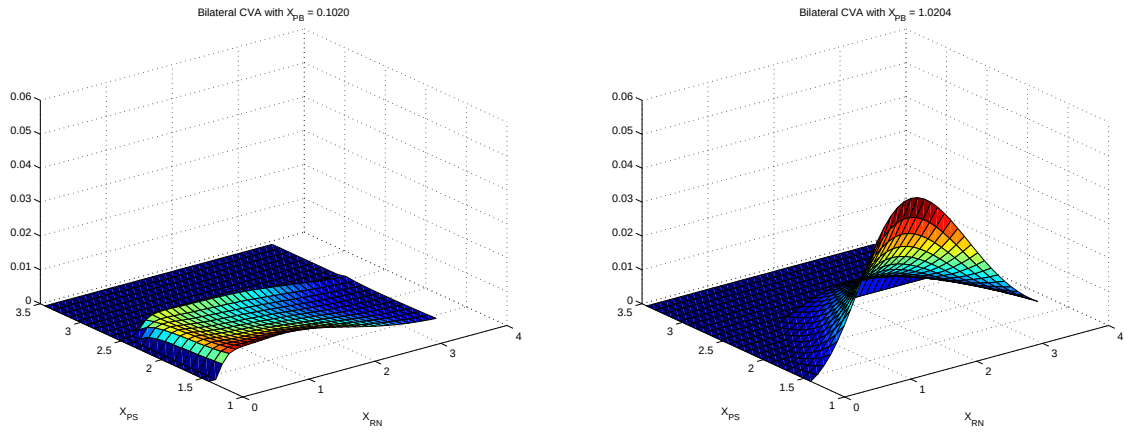


(c) (d)



(e) (f)

Figure 3.5: Joint survival probability for 3 banks as a projection on the y-z plane for different values of x.



(a) (b)

(c) (d)

Figure 3.6: CVA as a projection on the z-x plane for different values of y. We plot on a truncated domain to better show the impact of the protection buyer on CVA. In (a) there is a small hump due numerical error (assets for PB are small, while for PS are large).

Chapter 4

Finite-difference splitting methods for the jump model

In this chapter, we focus on the numerical computation of survival probabilities and prices of credit products in the extended model with jumps, where closed-form expressions are no longer available. Due to the curse of dimensionality, PDE methods are not efficient for a growth of the dimension, therefore they are usually used for $N \leq 3$.

Our work is most similar to Itkin and Lipton (2015), who develop a finite difference method for the partial integro-differential equation (PIDE) where the integral term results from a fairly general correlated Lévy process in the jump component. By Strang splitting into the diffusion and jump operators, overall second order consistency in the timestep is obtained. Hereby, the multi-dimensional diffusion operator is itself split dimensionwise using the Hundsdorfer-Verwer (HV) scheme, and the jump operator is treated as a pseudo-differential operator, which allows efficient evaluation of the discretised operator by an iterative procedure. Stability of each of the steps is guaranteed under standard conditions.

Our approach is more straightforward in that we apply a modification of the HV scheme directly, where we treat the jump term in the same way as the cross-derivative term in the classical HV scheme. For the analysis we consider infinite meshes, i.e. ignoring the boundaries, such that the discrete operators are also infinite-dimensional. In our analysis we build heavily on the results in In't Hout and Welfert (2007) on stability of the PDE with cross-derivatives (but no integral term) and periodic boundary conditions on a finite mesh.

We show that the (unconditional) von Neumann stability of the scheme is not materially affected by the jump operator, as its contribution to the symbol of the scheme is of a lower order in the mesh size. For concreteness, we restrict ourselves to the model with negative exponential jumps described in Lipton (2016). This allows a simple recursive computation of the discrete jump operator and gives an explicit form of its eigenvalues. However, the analysis can in principle be extended to other jump size distributions. Recent papers by in't Hout and Toivanen (2018) and Boen and Hout (2019) improve our splitting scheme ADI scheme and stability analyses for a general jump process,

A survey of splitting methods in finance is found in Toivanen et al. (2016). These are roughly arranged in two groups: splitting by dimension (for multi-dimensional PDEs; such as in In't Hout and Welfert (2007)), and splitting by operator type (for PIDEs,

diffusion and jumps; such as in Andersen and Andreasen (2000)). Itkin and Lipton (2015) perform these two splittings successively as described above. To our knowledge, the present paper is the first to perform and analyse splitting into dimensions and jumps simultaneously.

The scheme is constructed to be second order consistent with the continuous integro-differential operator applied to smooth functions. However, the discontinuities in the data lead to empirically observed convergence of only first order in both space and time step. To address this, we apply a spatial smoothing technique discussed in Pooley et al. (2003) for discontinuous option payoffs in the one-dimensional setting, and a change of the time variable to square-root time (see Reisinger and Whitley (2014)), equivalent to a quadratically refined time mesh close to maturity, in order to restore second order convergence. We emphasise that the presence of discontinuous initial data is essential to the nature of P(I)DE models of credit risk. Hence this approach improves on previous works in a key aspect of the numerical solution.

Similar to Itkin and Lipton (2017), we restrict the analysis to the two-dimensional case, but there is no fundamental problem in extending the method to multiple dimensions. However, due to the curse of dimensionality, for more than three-dimensional problems, standard finite-difference methods are computationally too expensive. The two-dimensional case already allows us to investigate various important model characteristics, such as joint and marginal survival probabilities, prices of credit derivatives, Credit and Debt Value Adjustments, and specifically the impact of mutual obligations.

Moreover, we provide a calibration of the model with negative exponential jumps to market data, and for this calibrated model assess the impact of mutual obligations on survival probabilities.

In Section 4.1, we consider the one-dimensional model, we show how to solve the pricing equation using a finite difference scheme and the Laplace transform method (both methods are described in Lipton and Sepp (2013)). In Section 4.2, we consider the model with two banks, we develop a numerical finite difference method, which is a modification of an ADI scheme, show its stability and convergence rate as well as how to improve the convergence rate to deal with discontinuous boundary and initial conditions. For the model with $N \geq 3$, a modification of the same scheme can be applied, therefore we restrict ourselves to the two-dimensional problem in this chapter. In Section 4.3, we consider some financial applications of the model. We compute joint and marginal survival probabilities, CDS and FTD spreads, CVA and DVA. Moreover, we show our calibration results to the market data.

4.1 Numerical methods for a one-dimensional problem

For a one-dimensional problem, (2.58)–(2.60) can be rewritten as

$$\frac{\partial V}{\partial t} + \xi \frac{\partial V}{\partial x} + \frac{1}{2}\frac{\partial^2 V}{\partial x^2} + \lambda \mathcal{J} V - \lambda V = \chi(t,x), \tag{4.1}$$

$$V(t,0) = \phi_0(t), \quad V(t,x) \underset{x\to+\infty}{\longrightarrow} \phi_\infty(t), \tag{4.2}$$

$$V(T,x) = \psi(x), \tag{4.3}$$

where $\chi, \phi_0, \phi_\infty, \psi$ are given functions as previously.

In order to deal with a forward equation instead of a backward equation, we change the time variable to $\tau = T - t$. Then, (4.1)–(4.3) transform to

$$\frac{\partial V}{\partial \tau} = \xi \frac{\partial V}{\partial x} + \frac{1}{2}\frac{\partial^2 V}{\partial x^2} + \lambda \mathcal{J} V - \lambda V - \chi(\tau, x), \tag{4.4}$$

$$V(\tau, 0) = \phi_0(\tau), \quad V(\tau, x) \underset{x \to +\infty}{\longrightarrow} \phi_\infty(\tau), \tag{4.5}$$

$$V(0, x) = \psi(x). \tag{4.6}$$

In this section, we show how to solve (4.1)–(4.3) numerically. Following Lipton and Sepp (2013), in Section 4.1.1 we show how to solve it using the Laplace transform, and using a finite–difference method in Section 4.1.2.

4.1.1 Laplace transform method

We will find the solution using the Green's function as we described in Chapter 3 . Writing (3.2) for (4.1)–(4.3), we get

$$\frac{\partial G}{\partial \tau} + \xi \frac{\partial G}{\partial x} - \frac{1}{2}\frac{\partial^2 G}{\partial x^2} - \lambda \mathcal{J} G + \lambda G = 0, \tag{4.7}$$

$$G(\tau, x', 0) = 0, \quad G(\tau, x', x) \underset{x \to +\infty}{\longrightarrow} 0, \tag{4.8}$$

$$G(0, x', x) = \delta(x' - x). \tag{4.9}$$

Once the Green's function is found, $V(t, x)$ can be easily computed using (3.3). For the one-dimensional problem, this simplifies to

$$V(t,x) = \int_0^{+\infty} \psi(x) G(T - t, x', x)\, \mathrm{d}x' + \frac{1}{2}\int_t^T \phi_0(t) G_x(t' - t, 0, x)\, \mathrm{d}t' - \int_t^T \int_0^\infty \chi(t', x') G(t' - t, x', x)\, \mathrm{d}x' \mathrm{d}t'. \tag{4.10}$$

Applying the Laplace transform in $\tau, G(\tau, x', x) \to \hat{G}(p, x', x)$ for (4.7)–(4.9), we get

$$p\hat{G}(p, x, x') + \xi \frac{\partial \hat{G}}{\partial x} - \frac{1}{2}\frac{\partial^2 \hat{G}}{\partial x^2} - \lambda \mathcal{J} G + \lambda G = \delta(x' - x). \tag{4.11}$$

The characteristic equation for the last equation has the form

$$\frac{1}{2}\psi^2 - \xi\psi - (\lambda + p) + \frac{\lambda \varsigma}{\varsigma - \psi} = 0. \tag{4.12}$$

This equation has three roots $\psi_j, 1 \le j \le 3$, and the solution has the form

$$\hat{G}(p, x, x') = \begin{cases} C_3 \mathrm{e}^{\psi_3 (x - x')}, & x \ge x', \\ \displaystyle\sum_{j=1}^{3} D_j \mathrm{e}^{\psi_j (x - x')}, & x < x', \end{cases} \tag{4.13}$$

where

$$D_1 = -\frac{2(\varsigma - \psi_1)}{(\psi_1 - \psi_2)(\psi_1 - \psi_3)}, \tag{4.14}$$

$$D_2 = -\frac{2(\varsigma - \psi_2)}{(\psi_2 - \psi_1)(\psi_2 - \psi_3)}, \tag{4.15}$$

$$D_3 = -\mathrm{e}^{-(\psi_1 - \psi_3)x} D_1 - \mathrm{e}^{-(\psi_2 - \psi_3)x} D_2, \tag{4.16}$$

$$C_3 = D_1 + D_2 + D_3. \tag{4.17}$$

The inverse Laplace transform can be found numerically, for example, using the Gaver-Stehfest algorithm (Abate et al. (2000)).

Using this method, the Green's function can be expressed as

$$G(\tau, x', x) \approx \frac{\ln 2}{t} \sum_{k=1}^{2m} \omega_k \hat{G}\left(\frac{k \ln 2}{\tau}, x', x\right),$$

where

$$\omega_k = (-1)^{m+k} \sum_{j=\lfloor (k+1)/2 \rfloor}^{\min(k,m)} \frac{j^{m+1}}{m} C_m^j C_{2j}^j C_{k-j}^j.$$

As an example of using the Gaver-Stehfest algorithm in finance, one can refer to Lipton and Sepp (2011), where the authors used it for the calibration of a local volatility surface.

4.1.2 Finite–difference method

We consider the same grid for integral and differential part of the equation

$$0 = x^0 < x^1 < \ldots < x^m, \tag{4.18}$$

where x^m is a large positive number.

The grid is non-uniform, and is chosen such that relatively many points lie near the default boundaries for better precision. We use a method similar to Itkin and Carr (2011) to construct the grid.

4.1.2.1 Discretization of the integral part of the PIDE

In this section, we shall show how to deal with the integral part of the PIDE, and develop an iterative algorithm for the fast computation of the integral operator on the grid.

The first approach is to deal with the integral operators directly. After the approximation of the integral, we get (Lipton and Sepp (2013))

$$\mathcal{J}V(x+h) = \mathrm{e}^{-\varsigma h} \mathcal{J}V(x) + \omega_0(\varsigma, h)V(x) + \omega(\varsigma, h)V(x) + O(h^3), \tag{4.19}$$

where

$$\omega_0(\varsigma, h) = \frac{1 - (1 + \varsigma h)\mathrm{e}^{-\varsigma h}}{\varsigma h}, \quad \omega_1(\varsigma, h) = \frac{-1 + \varsigma h + \mathrm{e}^{-\varsigma h}}{\varsigma h}.$$

Then, we can write recurrence formulas for computing the integral operator on the grid. Denote J^i the corresponding approximations of $\mathcal{J}V(x^i)$ on the grid. Applying (4.19), we get

$$J^{i+1} = \mathrm{e}^{-\varsigma h_{i+1}} J^i + \omega_0(\varsigma, h_{i+1})V(x^i) + \omega_1(\varsigma, h_{i+1})V(x^{i+1}), \tag{4.20}$$

where $h_{i+1} = x^{i+1} - x^i$.

For an alternative method, we rewrite the integral operator as a differential equation

$$\frac{\partial}{\partial x}\left(\mathcal{J}V(x)\mathrm{e}^{\varsigma x}\right) = \varsigma V(x)\mathrm{e}^{\varsigma}. \tag{4.21}$$

Then, we apply the Adams-Moulton method of second order which gives us third order of accuracy locally (Butcher (2008))

$$J^{i+1} = \mathrm{e}^{-\varsigma h_{i+1}} J^i + \frac{1}{2}h_{i+1}\mathrm{e}^{-\varsigma h_{i+1}}\varsigma V(x^i) + \frac{1}{2}h_{i+1}\varsigma V(x^{i+1}), \tag{4.22}$$

where $h_{i+1} = x^{i+1} - x^i$, and is equivalent to the trapezoidal rule.

We can rewrite (4.22) in the same notation as (4.20) by defining

$$\omega_0(\varsigma, h) = \frac{1}{2}h\mathrm{e}^{-\varsigma h}\varsigma, \quad \omega_1(\varsigma, h) = \frac{1}{2}h\varsigma.$$

So,

$$J^{i+1} = \mathrm{e}^{-\varsigma h_{i+1}} J^i + \omega_0(\varsigma, h_{i+1})V(x^i) + \omega_1(\varsigma, h_{i+1})V(x^{i+1}). \tag{4.23}$$

As a result we get explicit recursive formulas for approximations of $\mathcal{J}V$ that can be computed for all grid points via $O(m)$ operations. Both methods give the same order of accuracy.

4.1.2.2 Discretization of the differential part of the PIDE

Now consider the approximation of derivatives in the differential operator on a non-uniform grid. We use the standard derivative approximation (Kluge (2002), In't Hout and Foulon (2010)). For the first derivative over each variable consider right-sided, central, and left-sided schemes. So, for the derivative over x we have:

$$\frac{\partial V}{\partial x}(x) \approx \alpha_{i,-2}V(x^{i-2}) + \alpha_{i,-1}V(x^{i-1}) + \alpha_{i,0}V(x^i), \tag{4.24}$$

$$\frac{\partial V}{\partial x}(x) \approx \beta_{i,-1}V(x^{i-1}) + \beta_{i,0}V(x^i) + \beta_{i,1}V(x^{i+1}), \tag{4.25}$$

$$\frac{\partial V}{\partial x}(x) \approx \gamma_{i,0}V(x^i) + \gamma_{i,1}V(x^{i+1}) + \gamma_{i,2}V(x^{i+2}), \tag{4.26}$$

with coefficients

$$\alpha_{i,-2} = \frac{\Delta x^i}{\Delta x^{i-1}(\Delta x^{i-1} + \Delta x^i)}, \quad \alpha_{i,-1} = \frac{-\Delta x^{i-1} - \Delta x^i}{\Delta x^{i-1}\Delta x^i}, \quad \alpha_{i,0} = \frac{\Delta x^{i-1} + 2\Delta x^i}{\Delta x^i(\Delta x^{i-1} + \Delta x^i)},$$

$$\beta_{i,-1} = \frac{-\Delta x^{i+1}}{\Delta x^i(\Delta x^i + \Delta x^{i+1})}, \quad \beta_{i,0} = \frac{\Delta x^{i+1} - \Delta x^i}{\Delta x^i\Delta x^{i+1}}, \quad \beta_{i,1} = \frac{\Delta x^i}{\Delta x^{i+1}(\Delta x^i + \Delta x^{i+1})},$$

$$\gamma_{i,0} = \frac{-2\Delta x^{i+1} - \Delta x^{i+2}}{\Delta x^{i+1}(\Delta x^{i+1} + \Delta x^{i+2})}, \quad \gamma_{i,1} = \frac{\Delta x^{i+1} + \Delta x^{i+2}}{\Delta x^{i+1}\Delta x^{i+2}}, \quad \gamma_{i,2} = \frac{-\Delta x^{i+1}}{\Delta x^{i+2}(\Delta x^{i+1} + \Delta x^{i+2})}.$$

For the boundaries at 0 we use the schemes (4.24), for the right boundaries at x^m, we use the schemes (4.26), and for other points we use the central schemes (4.25).

To approximate the second derivative we use the central scheme:

$$\frac{\partial^2 V}{\partial x^2}(x^i) \approx \delta_{i,-1} V(x^{i-1}) + \delta_{i,0} V(x^i) + \delta_{i,1} V(x^{i+1}), \tag{4.27}$$

with coefficients

$$\delta_{i,-1} = \frac{2}{\Delta x^i (\Delta x^i + \Delta x^{i+1})}, \quad \delta_{i,0} = \frac{-2}{\Delta x^i \Delta x^{i+1}}, \quad \delta_{i,1} = \frac{2}{\Delta x^{i+1} (\Delta x^i + \Delta x^{i+1})},$$

As a result, we can approximate the differential operator $\mathcal{D}V$ by a discrete operator DV.

4.1.2.3 Time discretization

Consider a uniform time mesh with time step Δt : $t_n = n\Delta t$ for $n = 0, \ldots, N$. For simplicity, we use an Euler type method in this section, which guarantees only first order of convergence in time. But, in principle, the scheme can be improved to second order using predictor-corrector methods. The idea is to treat the integral part of the equation explicitly, and the diffusion part of the equation implicitly by inverting the tridiagonal matrix. Denote $V_0 = V(t_0, x)$ and $\chi_i = \chi(t_i, x)$. Then,

$$U = V_{i-1} + \Delta t \left(\lambda J - \lambda V_{i-1} - \chi_i \right), \tag{4.28}$$

$$V_i = (I - \Delta t D)^{-1} U. \tag{4.29}$$

The overall complexity of the method is $O(Nm)$ operations, since the computation of the integral operator and inversion of the tridiagonal matrix can be done in $O(m)$ operations. The method guarantees second order of convergence in space and first order of convergence in time variables. The latter can be further improved by applying more advanced methods for time-stepping.

4.2 Numerical method for a multidimensional problem

We shall solve the PIDE (2.58)–(2.60) numerically with an ADI method. This scheme is a modification of Lipton and Sepp (2013).We improve it to accommodate the integral part (jumps) in Hundsdorfer–Verwer (HV) scheme, which gives us second order of convergence in both time and step variable. We also show that the scheme is unconditionally stable.

In order to deal with a forward equation instead of a backward equation, we change the time variable to $\tau = T - t$

$$\begin{aligned}
&\frac{\partial V}{\partial \tau} = \mathcal{L} V(\tau, x_1, x_2) - \chi(\tau, x_1, x_2), \\
&V(\tau, x_1, 0) = \phi_{0,1}(\tau, x_1), \quad V(\tau, 0, x_2) = \phi_{0,2}(\tau, x_2), \\
&V(\tau, x_1, x_2) \underset{x_2 \to +\infty}{\longrightarrow} \phi_{\infty,1}(\tau, x_1), \quad V(\tau, x_2, x_2) \underset{x_1 \to +\infty}{\longrightarrow} \phi_{\infty,2}(\tau, x_2), \\
&V(0, x_1, x_2) = \psi(x_1, x_2),
\end{aligned} \tag{4.30}$$

We consider the same grid for integral and differential part of the equation

$$\begin{aligned} 0 = x_1^0 < x_1^1 < \ldots < x_1^{m_1}, \\ 0 = x_2^0 < x_2^1 < \ldots < x_2^{m_2}, \end{aligned} \tag{4.31}$$

where $x_1^{m_1}$ and $x_2^{m_2}$ are large positive numbers.

The grid is non-uniform, and is chosen such that relatively many points lie near the default boundaries for better precision. We use a method similar to Itkin and Carr (2011) to construct the grid.

In the next two sections, we consider a multiple-dimensional extension of the methods from the previous section.

4.2.1 Discretization of the integral part of the PIDE

In this section, we shall show how to deal with the integral part of the PIDE, and develop an iterative algorithm for the fast computation of the integral operator on the grid. To this end, we outline the scheme from Lipton and Sepp (2013) and then give a new method.

As previously,

$$\mathcal{J}_1 V(x_1 + h, x_2) = \mathrm{e}^{-\varsigma_1 h} \mathcal{J}_1 V(x_1, x_2) + \omega_0(\varsigma_1, h) V(x_1, x_2) + \omega_1(\varsigma_1, h) V(x_1 + h, x_2) + O(h^3), \tag{4.32}$$

$$\mathcal{J}_2 V(x_1, x_2 + h) = \mathrm{e}^{-\varsigma_2 h} \mathcal{J}_2 V(x_1, x_2) + \omega_0(\varsigma_2, h) V(x_1, x_2) + \omega_1(\varsigma_2, h) V(x_1, x_2 + h) + O(h^3), \tag{4.33}$$

where

$$\omega_0(\varsigma, h) = \frac{1 - (1 + \varsigma h)\mathrm{e}^{-\varsigma h}}{\varsigma h}, \quad \omega_1(\varsigma, h) = \frac{-1 + \varsigma h + \mathrm{e}^{-\varsigma h}}{\varsigma h}.$$

We can also approximate $\mathcal{J}_{12} V = \mathcal{J}_1 \mathcal{J}_2 V$ by applying above approximations for $\mathcal{J}_1$ and $\mathcal{J}_2$ consecutively.

Consider the grid

$$\begin{aligned} 0 = x_1^0 < x_1^1 < \ldots < x_1^{m_1}, \\ 0 = x_2^0 < x_2^1 < \ldots < x_2^{m_2}, \end{aligned} \tag{4.34}$$

where $x_1^{m_1}$ and $x_2^{m_2}$ are large positive numbers.

Then, we can write recurrence formulas for computing the integral operator on the grid. Denote $J_1^{i,j}$, $J_2^{i,j}$, $J_{12}^{i,j}$ the corresponding approximations of $\mathcal{J}_1 V(x_1^i, x_2^j)$, $\mathcal{J}_2 V(x_1^i, x_2^j)$, $\mathcal{J}_{12} V(x_1^i, x_2^j)$ on the grid. Applying (4.32) and (4.33) we get

$$J_1^{i+1,j} = \mathrm{e}^{-\varsigma_1 h_{i+1}^1} J_1^{i,j} + \omega_0(\varsigma_1, h_{i+1}^1) V(x_1^i, x_2^j) + \omega_1(\varsigma_1, h_{i+1}^1) V(x_1^{i+1}, x_2^j), \tag{4.35}$$

$$J_2^{i,j+1} = \mathrm{e}^{-\varsigma_2 h_{j+1}^2} J_2^{i,j} + \omega_0(\varsigma_2, h_{j+1}^2) V(x_1^i, x_2^j) + \omega_1(\varsigma_2, h_{j+1}^2) V(x_1^i, x_2^{j+1}), \tag{4.36}$$

where $h_{i+1}^1 = x_1^{i+1} - x_1^i, h_{j+1}^2 = x_2^{j+1} - x_2^j$.

For an alternative method, we rewrite the integral operator as a differential equation

$$\frac{\partial}{\partial x_1}\left(\mathcal{J}_1 V(x_1,x_2)\mathrm{e}^{\varsigma_1 x_1}\right) = \varsigma_1 V(x_1,x_2)\mathrm{e}^{\varsigma_1 x_1}, \tag{4.37}$$

$$\frac{\partial}{\partial x_2}\left(\mathcal{J}_2 V(x_1,x_2)\mathrm{e}^{\varsigma_2 x_2}\right) = \varsigma_2 V(x_1,x_2)\mathrm{e}^{\varsigma_2 x_2}, \tag{4.38}$$

$$\frac{\partial^2}{\partial x_1 \partial x_2}\left(\mathcal{J}_{12} V(x_1,x_2)\mathrm{e}^{\varsigma_1 x_1+\varsigma_2 x_2}\right) = \varsigma_1\varsigma_2 V(x_1,x_2)\mathrm{e}^{\varsigma_1 x_1+\varsigma_2 x_2}. \tag{4.39}$$

Then, we apply the Adams-Moulton method of second order which gives us third order of accuracy locally (Butcher (2008))

$$J_1^{i+1,j} = \mathrm{e}^{-\varsigma_1 h_{i+1}^1} J_1^{i,j} + \frac{1}{2} h_{i+1}^1 \mathrm{e}^{-\varsigma_1 h_{i+1}^1}\varsigma_1 V(x_1^i,x_2^j) + \frac{1}{2} h_{i+1}^1 \varsigma_1 V(x_1^{i+1},x_2^j), \tag{4.40}$$

$$J_2^{i,j+1} = \mathrm{e}^{-\varsigma_2 h_{j+1}^2} J_2^{i,j} + \frac{1}{2} h_{j+1}^2 \mathrm{e}^{-\varsigma_2 h_{j+1}^2}\varsigma_2 V(x_1^i,x_2^j) + \frac{1}{2} h_{j+1}^2 \varsigma_2 V(x_1^i,x_2^{j+1}), \tag{4.41}$$

where $h_{i+1}^1 = x_1^{i+1} - x_1^i, h_{j+1}^2 = x_2^{j+1} - x_2^j$, and is equivalent to the trapezoidal rule.

We can rewrite (4.40)–(4.41) in the same notation as (4.35)–(4.36) by defining

$$\omega_0(\varsigma,h) = \frac{1}{2} h \mathrm{e}^{-\varsigma h}\varsigma, \quad \omega_1(\varsigma,h) = \frac{1}{2} h\varsigma.$$

So,

$$J_1^{i+1,j} = \mathrm{e}^{-\varsigma_1 h_{i+1}^1} J_1^{i,j} + \omega_0(\varsigma_1,h_{i+1}^1) V(x_1^i,x_2^j) + \omega_1(\varsigma_1,h_{i+1}^1) V(x_1^{i+1},x_2^j),$$
$$J_2^{i,j+1} = \mathrm{e}^{-\varsigma_2 h_{j+1}^2} J_2^{i,j} + \omega_0(\varsigma_2,h_{j+1}^2) V(x_1^i,x_2^j) + \omega_1(\varsigma_2,h_{j+1}^2) V(x_1^i,x_2^{j+1}).$$

As a result we get explicit recursive formulas for approximations of $\mathcal{J}_1 V$ and $\mathcal{J}_2 V$ that can be computed for all grid points via $O(m_1 m_2)$ operations. Both methods give the same order of accuracy. As was discussed above, in order to compute the approximation of $\mathcal{J}_{12} V$ we can apply consecutively the approximations of $\mathcal{J}_2 V$ and $\mathcal{J}_1(\mathcal{J}_2 V)$. So, we have the two-step procedure:

$$I_{12}^{i+1,j} = \mathrm{e}^{-\varsigma_1 h_{i+1}^1} I_{12}^{i,j} + \omega_0(\varsigma_1,h_{i+1}^1) V(x_1^i,x_2^j) + \omega_1(\varsigma_1,h_{i+1}^1) V(x_1^{i+1},x_2^j), \tag{4.42}$$

and

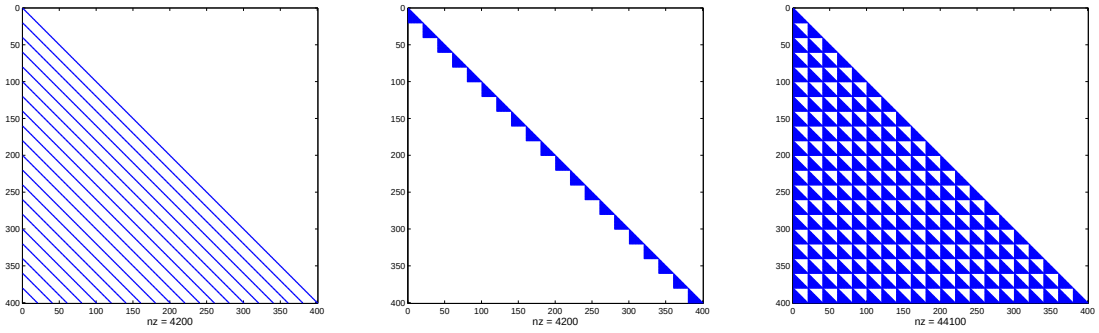
$$J_{12}^{i,j+1} = \mathrm{e}^{-\varsigma_2 h_{j+1}^2} J_{12}^{i,j} + \omega_0(\varsigma_2,h_{j+1}^2) I_{12}^{i,j} + \omega_1(\varsigma_2,h_{j+1}^2) I_{12}^{i,j+1}. \tag{4.43}$$

Using this two-step procedure, we can also compute an approximation of $\mathcal{J}_{12} V$ on the grid in complexity $O(m_1 m_2)$.

We shall subsequently analyze the stability of the second method and use it in the numerical tests. The results for the first method would be very similar.

For the implementation, computing and storing a matrix representation of the jump operator is not necessary, since the operator can be computed iteratively as described above, but we shall use matrix notation for the analysis. We henceforth denote J_1, J_2, and J_{12} the matrices of the discretized jump operators. From (4.35)–(4.36) we can find that the matrices J_1 and J_2 are lower-triangular with diagonal elements $w_1 = \omega_1(\varsigma_1,h_1)$ and $w_2 = \omega_1(\varsigma_2,h_2)$. Then, $J_{12} = J_1 J_2$ is also a lower-triangular matrix with diagonal

elements $w_1 w_2$. To illustrate, in Figure 4.1 we plot the sparsity patterns in J_1, J_2, and J_{12}.



(a) J_1. (b) J_2. (c) J_{12}.

Figure 4.1: Sparsity pattern of J_1, J_2, and J_{12}. Here, $m_1 = m_2 = 20$ and nz is the number of non-zero elements of the matrices.

4.2.2 Discretization of the differential part of the PIDE

Now consider the approximation of derivatives in the differential operator on a non-uniform grid. As before, we use the approximation methods from (Kluge (2002), In't Hout and Foulon (2010)). For the first derivative over each variable consider right-sided, central, and left-sided schemes. So, for the derivative over x_1 we have:

$$\frac{\partial V}{\partial x_1}(x_1^i, x_2^j) \approx \alpha^1_{i,-2} V(x_1^{i-2}, x_2^j) + \alpha^1_{i,-1} V(x_1^{i-1}, x_2^j) + \alpha^1_{i,0} V(x_1^i, x_2^j), \tag{4.44}$$

$$\frac{\partial V}{\partial x_1}(x_1^i, x_2^j) \approx \beta^1_{i,-1} V(x_1^{i-1}, x_2^j) + \beta^1_{i,0} V(x_1^i, x_2^j) + \beta^1_{i,1} V(x_1^{i+1}, x_2^j), \tag{4.45}$$

$$\frac{\partial V}{\partial x_1}(x_1^i, x_2^j) \approx \gamma^1_{i,0} V(x_1^i, x_2^j) + \gamma^1_{i,1} V(x_1^{i+1}, x_2^j) + \gamma^1_{i,2} V(x_1^{i+2}, x_2^j), \tag{4.46}$$

while for derivative over x_2 we have:

$$\frac{\partial V}{\partial x_2}(x_1^i, x_2^j) \approx \alpha^2_{j,-2} V(x_1^i, x_2^{j-2}) + \alpha^2_{j,-1} V(x_1^i, x_2^{j-1}) + \alpha^2_{j,0} V(x_1^i, x_2^j), \tag{4.47}$$

$$\frac{\partial V}{\partial x_2}(x_1^i, x_2^j) \approx \beta^2_{j,-1} V(x_1^i, x_2^{j-1}) + \beta^2_{j,0} V(x_1^i, x_2^j) + \beta^2_{j,1} V(x_1^j, x_2^{j+1}), \tag{4.48}$$

$$\frac{\partial V}{\partial x_2}(x_1^i, x_2^j) \approx \gamma^2_{j,0} V(x_1^i, x_2^j) + \gamma^2_{j,1} V(x_1^i, x_2^{j+1}, x_2^j) + \gamma^2_{j,2} V(x_1^i, x_2^{j+2}), \tag{4.49}$$

with coefficients

$$\alpha^k_{i,-2} = \frac{\Delta x_k^i}{\Delta x_k^{i-1}(\Delta x_k^{i-1} + \Delta x_k^i)}, \quad \alpha^k_{i,-1} = \frac{-\Delta x_k^{i-1} - \Delta x_k^i}{\Delta x_k^{i-1} \Delta x_k^i}, \quad \alpha^k_{i,0} = \frac{\Delta x_k^{i-1} + 2\Delta x_k^i}{\Delta x_k^i (\Delta x_k^{i-1} + \Delta x_k^i)},$$

$$\beta^k_{i,-1} = \frac{-\Delta x_k^{i+1}}{\Delta x_k^i (\Delta x_k^i + \Delta x_k^{i+1})}, \quad \beta^k_{i,0} = \frac{\Delta x_k^{i+1} - \Delta x_k^i}{\Delta x_k^i \Delta x_k^{i+1}}, \quad \beta^k_{i,1} = \frac{\Delta x_k^i}{\Delta x_k^{i+1} (\Delta x_k^i + \Delta x_k^{i+1})},$$

$$\gamma^k_{i,0} = \frac{-2\Delta x_k^{i+1} - \Delta x_k^{i+2}}{\Delta x_k^{i+1} (\Delta x_k^{i+1} + \Delta x_k^{i+2})}, \quad \gamma^k_{i,1} = \frac{\Delta x_k^{i+1} + \Delta x_k^{i+2}}{\Delta x_k^{i+1} \Delta x_k^{i+2}}, \quad \gamma^k_{i,2} = \frac{-\Delta x_k^{i+1}}{\Delta x_k^{i+2} (\Delta x_k^{i+1} + \Delta x_k^{i+2})}.$$

For the boundaries at 0 we use the schemes (4.44) and (4.47), for the right boundaries at $x_1^{m_1}$ and $x_2^{m_2}$ we use the schemes (4.46) and (4.49), and for other points we use the central schemes (4.45) and (4.48).

To approximate the second derivative we use the central scheme:

$$\frac{\partial^2 V}{\partial x_1^2}(x_1^i, x_2^j) \approx \delta_{i,-1}^1 V(x_1^{i-1}, x_2^j) + \delta_{i,0}^1 V(x_1^i, x_2^j) + \delta_{i,1}^1 V(x_1^{i+1}, x_2^j), \tag{4.50}$$

$$\frac{\partial^2 V}{\partial x_2^2}(x_1^i, x_2^j) \approx \delta_{j,-1}^2 V(x_1^i, x_2^{j-1}) + \delta_{j,0}^2 V(x_1^i, x_2^j) + \delta_{j,1}^2 V(x_1^j, x_2^{j+1}), \tag{4.51}$$

with coefficients

$$\delta_{i,-1}^k = \frac{2}{\Delta x_k^i(\Delta x_k^i + \Delta x_k^{i+1})}, \quad \delta_{i,0}^k = \frac{-2}{\Delta x_k^i \Delta x_k^{i+1}}, \quad \delta_{i,1}^k = \frac{2}{\Delta x_k^{i+1}(\Delta x_k^i + \Delta x_k^{i+1})},$$

and to approximate the second mixed derivative we use the scheme:

$$\frac{\partial^2 V}{\partial x_1 \partial x_2}(x_1^i, x_2^j) \approx \sum_{k,l=-1}^{1} \beta_{i,k}^1 \beta_{j,l}^2 V(x_1^{i+k}, x_2^{j+l}). \tag{4.52}$$

As a result, we can approximate the differential operator $\mathcal{D}V$ by a discrete operator

$$DV = D_1 V + D_2 V + D_{12} V, \tag{4.53}$$

where $D_1 V$ contains the discretized derivatives over x_1 defined in (4.44)–(4.46) and (4.50), $D_2 V$ contains the discretized derivatives over x_2 defined in (4.47)–(4.49) and (4.51), and $D_{12} V$ contains the discretized mixed derivative defined in (4.52).

By straightforward but lengthy Taylor expansion of the expression in (4.44)–(4.52), the scheme (4.57) has second order truncation error in variables x_1 and x_2 for meshes which are either uniform or smooth transformations of such meshes, as we shall consider later.

4.2.3 Time discretization: ADI scheme

After discretization over (x_1, x_2) we can rewrite the PIDE (4.30) as a system of ordinary (linear) differential equations. Consider the vector $U(t) \in \mathbb{R}^{m_1 m_2 \times 1}$ whose elements correspond to $V(t, x_1^i, x_2^j)$. Then

$$\begin{aligned} U'(t) &= \tilde{A}U(t) + b(t), \\ U(0) &= U_0, \end{aligned} \tag{4.54}$$

where $\tilde{A} = D_1 + D_2 + D_{12} + \lambda_1 J_1 + \lambda_2 J_2 + \lambda_{12} J_{12} - (\lambda_1 + \lambda_2 + \lambda_{12})I$, and $b(t)$ is determined from boundary conditions and the right-hand side.

To solve this system, we apply an ADI scheme for the time discretization. Consider, for simplicity, a uniform time mesh with time step $\Delta t : t_n = n\Delta t, n = 0, \ldots, N-1$.

We decompose the matrix $\tilde{A}$ into three matrices, $\tilde{A} = \tilde{A}_0 + \tilde{A}_1 + \tilde{A}_2$, where

$$\begin{aligned}
\tilde{A}_0 &= D_{12} + \lambda_1 J_1 + \lambda_2 J_2 + \lambda_{12} J_{12}, \\
\tilde{A}_1 &= D_1 - \left(\lambda_1 + \frac{\lambda_{12}}{2}\right) I, \\
\tilde{A}_2 &= D_2 - \left(\lambda_2 + \frac{\lambda_{12}}{2}\right) I,
\end{aligned}$$

and $b(t) = b_0(t) + b_1(t) + b_2(t)$, where $b_0(t)$ corresponds to the right-hand side and the FD discretization of the mixed derivatives on the boundary, $b_1(t)$ and $b_2(t)$ correspond to the FD discretization of the derivatives over x_1 and x_2 on the boundary.

Now we can apply a traditional ADI scheme with matrices $\tilde{A}_0$, $\tilde{A}_1$, and $\tilde{A}_2$. We choose the Hundsdorfer–Verwer (HV) scheme (Hundsdorfer and Verwer (2013)) in order to have second order accuracy in the time variable, and unconditional stability, as we shall prove below. For convenience, denote

$$F_j(t,x) = \tilde{A}_j x + b_j(t), \quad j = 0,1,2, \tag{4.55}$$

$$F(t,x) = (\tilde{A}_0 + \tilde{A}_1 + \tilde{A}_2)x + (b_0(t) + b_1(t) + b_2(t)), \tag{4.56}$$

and apply the Hundsdorfer–Verwer (HV) scheme:

$$\begin{cases}
Y_0 = U_{n-1} + \Delta t F(t_{n-1}, U_{n-1}), \\
Y_j = Y_{j-1} + \theta \Delta t (F_j(t_n, Y_j) - F_j(t_n, U_{n-1})), \quad j = 1,2, \\
\tilde{Y}_0 = Y_0 + \sigma \Delta t (F(t_n, Y_2) - F(t_{n-1}, U_{n-1})), \\
\tilde{Y}_j = \tilde{Y}_{j-1} + \theta \Delta t (F_j(t_n, \tilde{Y}_j - F_j(t_n, Y_2)), \quad j = 1,2, \\
U_n = \tilde{Y}_2.
\end{cases} \tag{4.57}$$

In this scheme, parts that contain F_1 and F_2 are treated implicitly. The matrix $\tilde{A}_1$ is tridiagonal and $\tilde{A}_2$ is block-tridiagonal and can be inverted via $O(m_1 m_2)$ operations. As a result, the overall complexity is $O(m_1 m_2)$ for a single time step or $O(N m_1 m_2)$ for the whole procedure.

Moreover, the scheme has second order of consistency in both (x_1, x_2) and t for any given θ and $\sigma = \frac{1}{2}$.

4.2.4 Stability analysis

In this section, we consider the PIDE (4.30) with zero boundary conditions at 0 in both directions and on a uniform grid, such that $F_j(t,x) = \tilde{A}_j x$ and

$$\begin{cases}
Y_0 = U_{n-1} + \Delta t \tilde{A} U_{n-1}, \\
Y_j = Y_{j-1} + \theta \Delta t (\tilde{A}_j Y_j - \tilde{A}_j U_{n-1}), \quad j = 1,2, \\
\tilde{Y}_0 = Y_0 + \sigma \Delta t (\tilde{A} Y_2 - \tilde{A} U_{n-1}), \\
\tilde{Y}_j = \tilde{Y}_{j-1} + \theta \Delta t (\tilde{A}_j \tilde{Y}_j - \tilde{A}_j Y_2), \quad j = 1,2 \\
U_n = \tilde{Y}_2.
\end{cases} \tag{4.58}$$

For convenience, we denote by $F : U_n = FU_{n-1}$.

We further consider the PDE on $\mathbb{R}^2$, i.e., without default boundaries. Hence, we assume that diffusion and jump operators are discretized on an infinite, uniform mesh $\{(j_1h_1, j_2h_2), (j_1, j_2) \in \mathbb{Z}^2\}$, such that, e.g. $D_1, D_2, D_{12}, J_1, J_2$ are infinite matrices. This is different to In't Hout and Welfert (2007), where finite matrices and periodic boundary conditions (without integral terms) are considered.

We use von Neumann stability analysis, as first introduced by Charney et al. (1950), by expanding the solution into a Fourier series. Hence, we shall show that the proposed scheme (4.58) is unconditionally stable, i.e. we will show that all eigenvalues of the operator F have moduli bounded by 1 plus an $O(\Delta t)$ term, where the corresponding eigenfunctions are given by $\exp(i\phi_1 j_1)\exp(i\phi_2 j_2)$, with ϕ_1 and ϕ_2 the wave numbers and j_1 and j_2 the grid coordinates.

In't Hout and Welfert (2007) show that when all matrices commute (as in the PDE case with periodic boundary conditions), the eigenvalues for F are given by

$$\begin{aligned} T(\tilde{z}_0, \tilde{z}_1, \tilde{z}_2) &= 1 + 2\frac{\tilde{z}_0 + \tilde{z}}{p} - \frac{\tilde{z}_0 + \tilde{z}}{p^2} + \sigma\frac{(\tilde{z}_0 + \tilde{z})^2}{p^2} \quad \text{with} \\ p &= (1 - \theta\tilde{z}_1)(1 - \theta\tilde{z}_2), \end{aligned} \tag{4.59}$$

where $\tilde{z}_j = \tilde{\mu}_j\Delta t$, where $\tilde{\mu}_j$ is an eigenvalue of $\tilde{A}_j$, $j = 0, 1, 2$, $\tilde{z} = \tilde{z}_1 + \tilde{z}_2$, $\theta \geq 0$.

The analysis is made slightly more complicated in our case through the presence of the jump operators. In the remainder of this section, we show that stability is still given under the same conditions on θ and σ as in the purely diffusive case. For the correspondence of notation with In't Hout and Welfert (2007), we denote $A = A_0 + A_1 + A_2$, where $A_0 = D_{12}, A_1 = D_1$, $A_2 = D_2$ and μ_0, μ_1, and μ_2 are the eigenvalues of the corresponding matrices. Similar to $\tilde{z}_0, \tilde{z}_1$, and $\tilde{z}_2$, we define scaled eigenvalues $z_0 = \mu_0\Delta t, z_1 = \mu_1\Delta t, z_2 = \mu_2\Delta t$.

We have the eigenvalues $\tilde{\mu}_j$ of $\tilde{A}_j$ given by

$$\tilde{\mu}_0 = \mu_0 + \lambda_1 w_1 + \lambda_2 w_2 + \lambda_{12} w_{12}, \tag{4.60}$$

$$\tilde{\mu}_1 = \mu_1 - \left(\lambda_1 + \frac{\lambda_{12}}{2}\right), \tag{4.61}$$

$$\tilde{\mu}_2 = \mu_2 - \left(\lambda_2 + \frac{\lambda_{12}}{2}\right), \tag{4.62}$$

where μ_j is an eigenvalue of A_j, and w_1, w_2, and w_{12} are eigenvalues of J_1, J_2, and J_{12}.

Denote $z = z_1 + z_2$, $s_1 = w_1\Delta t, s_2 = w_2\Delta t, s_{12} = w_{12}\Delta t$, where w_1, w_2, w_{12} are eigenvalues of J_1, J_2, J_{12} respectively, and $s_0 = \lambda_1 s_1 + \lambda_2 s_2 + \lambda_{12} s_{12}$.

Multiplying (4.60)–(4.62) by Δt, we have

$$\tilde{z}_0 = z_0 + s_0, \tag{4.63}$$

$$\tilde{z}_1 = z_1 - \left(\lambda_1 + \frac{\lambda_{12}}{2}\right)\Delta t, \tag{4.64}$$

$$\tilde{z}_2 = z_2 - \left(\lambda_2 + \frac{\lambda_{12}}{2}\right)\Delta t. \tag{4.65}$$

Theorem 4.1 (In't Hout and Welfert (2007), Theorem 3.2)**.** *Assume* $\Re(z_1) \le 0, \Re(z_2) \le 0$, $|z_0| \le 2\sqrt{\Re(z_1)\Re(z_2)}$, *where* z_0, z_1, *and* z_2 *are the eigenvalues of* A_0, A_1, *and* A_2, *and*

$$\frac{1}{2} \le \sigma \le \left(1 + \frac{\sqrt{2}}{2}\right)\theta.$$

Then,

$$|T(z_0, z_1, z_2)| \le 1,$$

and the Hundsdorfer–Verwer scheme (4.58) *is stable in the purely diffusive case.*

Lemma 4.1. *The scaled eigenvalues of* A_0, A_1, A_2, J_1, J_2, J_{12} *can be expressed as*

$$\begin{aligned}
z_0 &= -\rho b[\sin\phi_1 \sin\phi_2], && (4.66)\\
z_1 &= -a_1(1-\cos\phi_1) + i\xi_1 q_1 \sin\phi_1, && (4.67)\\
z_2 &= -a_2(1-\cos\phi_2) + i\xi_2 q_2 \sin\phi_2, && (4.68)\\
s_1 &= \Delta t\, \zeta_1 h_1 \left(\frac{1}{2} + \frac{\exp(-h_1\zeta_1 + i\phi_1)}{1-\exp(-h_1\zeta_1 + i\phi_1)}\right), && (4.69)\\
s_2 &= \Delta t\, \zeta_2 h_2 \left(\frac{1}{2} + \frac{\exp(-h_2\zeta_2 + i\phi_2)}{1-\exp(-h_2\zeta_2 + i\phi_2)}\right), && (4.70)\\
s_{12} &= s_1 s_2/\Delta t, && (4.71)
\end{aligned}$$

where

$$q_1 = \frac{\Delta t}{h_1}, \quad q_2 = \frac{\Delta t}{h_2}, \quad a_1 = \frac{\Delta t}{h_1^2}, \quad a_2 = \frac{\Delta t}{h_2^2}, \quad b = \frac{\Delta t}{h_1 h_2},$$

and $\phi_j \in [0, 2\pi]$ *for* $j = 1, 2$.

Moreover,

$$|z_0| \le 2\sqrt{\Re(z_1)\Re(z_2)}. \tag{4.72}$$

Proof. All six eigenvalues follow by insertion of the ansatz $U = \exp(i\phi_1 j_1)\exp(i\phi_2 j_2)$. For instance,

$$(J_1 U)(j_1, j_2) = \zeta_1 h_1 \left(\frac{1}{2}U(j_1, j_2) + \sum_{k=1}^{\infty} \exp(-\zeta_1 h_1 k) U(j_1 - k, j_2)\right),$$

and the result follows by using the special form of U and evaluating the geometric series.

Alternatively, the first three equations follow immediately from the eigenvalues for finite matrices (In't Hout and Welfert (2007), p.29), which are given by (4.66)–(4.68) where $\phi_j = 2l\pi/m_j$, $l = 1, \ldots, m_j$. In the infinite mesh case, the spectrum is the continuous limit and (4.72) still holds. □

Theorem 4.2. *Consider* $\frac{1}{2} \le \sigma \le \left(1 + \frac{\sqrt{2}}{2}\right)\theta$. *Then there exists* $c > 0$, *independent of* $\Delta t \le 1$, h_1 *and* h_2, *such that*

1. $$|T(\tilde{z}_0, \tilde{z}_1, \tilde{z}_2)| \le 1 + c\Delta t, \qquad \forall \phi_1, \phi_2 \in [0, 2\pi], \tag{4.73}$$

   i.e., the scheme is von Neumann stable;

2.

$$|U_n|_2 \le \mathrm{e}^{cn\Delta t}|U_0|_2, \qquad \forall n \ge 0, \tag{4.74}$$

for $|U_n|_2 = h_1 h_2 \left(\sum_{j_1,j_2=-\infty}^{\infty} |U_n(j_1,j_2)|^2\right)^{1/2}$*, i.e., the scheme is* l_2 *stable.*

Proof. First, we have that

$$|T(z_0,\tilde z_1,\tilde z_2)| = \left|1 + 2\frac{z_0+\tilde z}{p} - \frac{z_0+\tilde z}{p^2} + \sigma\frac{(z_0+\tilde z)^2}{p^2}\right| \le 1,$$

where as before $p = (1-\theta\tilde z_1)(1-\theta\tilde z_2)$ and $\tilde z = \tilde z_1 + \tilde z_2$. This follows from Theorem 4.1 because λ_1, λ_2 and λ_{12} are positive and therefore (4.72) is still satisfied with z_1 and z_2 replaced by $\tilde z_1$ and $\tilde z_2$.

We have

$$T(\tilde z_0,\tilde z_1,\tilde z_2) \;=\; T(z_0,\tilde z_1,\tilde z_2) + 2\frac{s_0}{p} - \frac{s_0}{p^2} + \sigma\frac{2s_0(z_0+\tilde z)+s_0^2}{p^2}.$$

A simple calculation shows that $|s_0| \le c_0\,\Delta t$ for a constant c_0 (independent of $\Delta t, h_1, h_2, \phi_1, \phi_2$; indeed, $c_0 = 2\lambda_1 + 2\lambda_2 + 4\lambda_{12}$ works for small enough h_1, h_2). Therefore, and because $|p| \ge 1$, $|z_0+\tilde z|/|p| \le c_1$ for a constant c_1,

$$\left|2\frac{s_0}{p} - \frac{s_0}{p^2} + \sigma\frac{2s_0(z_0+\tilde z)+s_0^2}{p^2}\right| \le c\Delta t,$$

for any $c \ge (3 + 2\sigma c_1 + c_0\sigma)c_0$. From this the first statement follows.

We can now deduce part 2 by a standard argument. For the discrete-continuous Fourier transform

$$l_2(\mathbb{Z}^2) \to L_2(-\pi,\pi)^2, \qquad U \to \widehat U, \qquad \widehat U(\phi_1,\phi_2) = h_1h_2\sum_{j,k\in\mathbb{Z}} U(j,k)\mathrm{e}^{-i(j\phi_1+k\phi_2)},$$

we have

$$\widehat U_{n+1}(\phi_1,\phi_2) = T(\tilde z_0,\tilde z_1,\tilde z_2)\,\widehat U_n(\phi_1,\phi_2), \qquad \forall n \ge 0.$$

Then, by Parseval,

$$\begin{aligned}
|U_n|_2^2 &= \frac{1}{4\pi^2}|\widehat U_n|^2 \\
&= \frac{1}{4\pi^2}\frac{1}{h_1^2h_2^2}\int_{-\pi}^{\pi}|\widehat U_n(\phi_1,\phi_2)|^2\,\mathrm{d}\phi_1\,\mathrm{d}\phi_2 \\
&\le \frac{1}{4\pi^2}\frac{1}{h_1^2h_2^2}\int_{-\pi}^{\pi}(1+c\Delta t)^{2n}|\widehat U_0(\phi_1,\phi_2)|^2\,\mathrm{d}\phi_1\,\mathrm{d}\phi_2 \\
&\le \mathrm{e}^{2cn\Delta t}\frac{1}{4\pi^2}\frac{1}{h_1^2h_2^2}\int_{-\pi}^{\pi}|\widehat U_0(\phi_1,\phi_2)|^2\,\mathrm{d}\phi_1\,\mathrm{d}\phi_2 \\
&= \mathrm{e}^{2cn\Delta t}|U_0|_2^2.
\end{aligned}$$

□

This (l_2-)stability result together with second order consistency implies (l_2-)convergence of second order for all solutions which are sufficiently smooth that the truncation error is defined and bounded. In our setting, where the initial condition is discontinuous, this is not given. Since the step function lies in the (l_2-)closure of smooth functions, convergence is guaranteed, but usually not of second order. We show this empirically in the next section and demonstrate how second order convergence can be restored practically.

4.2.5 Discontinuous boundary and terminal conditions

It is well documented (see, e.g. Pooley et al. (2003)) that the spatial convergence order of central finite difference schemes is generally reduced to one for discontinuous payoffs. Moreover, the time convergence order of the Crank-Nicolson scheme is reduced to one due to the lack of damping of high-frequency components of the error, and this behaviour is inherited by the HV scheme. We address these two issues in the following way.

First, we smooth the terminal condition by the method of local averaging from Pooley et al. (2003), i.e., instead of using nodal values of ϕ directly, we use the approximation

$$\phi(x_1^i, x_2^j) \approx \frac{1}{h_1 h_2} \int_{x_2^i - h_2/2}^{x_2^i + h_2/2} \int_{x_1^j - h_1/2}^{x_1^j + h_1/2} \phi(\xi_1, \xi_2)\, \mathrm{d}\xi_1 \mathrm{d}\xi_2.$$

For step functions with values of 0 and 1, this procedure attaches to each node the fraction of the area where the payoff is 1, in a cell of of size $h_1 \times h_2$ centred at this point.

We illustrate the convergence improvement on the example of joint survival probabilities. Other quantities show a similar behaviour. The model parameters in the following tests are the same as in the next section, specifically Table 4.1.

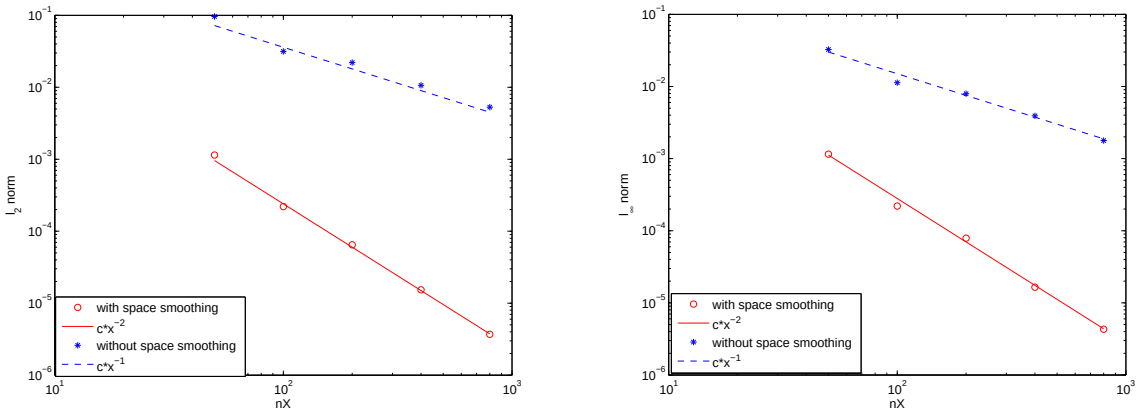
We choose $\sigma = \frac{1}{2}$ and $\theta = \frac{3}{4}$ in the HV scheme.

The observed convergence with and without this smoothing procedure is shown in Figure 4.2. We choose the l_2-norm for its closeness to the stability analysis – in the periodic case, Fourier analysis gives convergence results in l_2 – and the l_∞-norm for its relevance to the problem at hand, where we are interested in the solution pointwise. The behaviour in the l_1-norm is very similar.

Hereby, for a method of order $p \geq 1$ we estimate the error by extrapolation as

$$|Q^{nX}(x_1, x_2) - Q(x_1, x_2)| \approx \frac{1}{2^p - 1} |Q^{nX}(x_1, x_2) - Q^{nX/2}(x_1, x_2)|,$$

where Q is the exact solution, Q^{nX} the solution with nX mesh points, and the norms are computed by either taking the maximum over mesh points or numerical quadrature. Here, $nT = 1000$ is fixed.



(a) l_2-norm. (b) l_∞-norm.

Figure 4.2: Convergence analysis for l_2- and l_∞-norms of the error depending on the mesh size with fixed time-step.

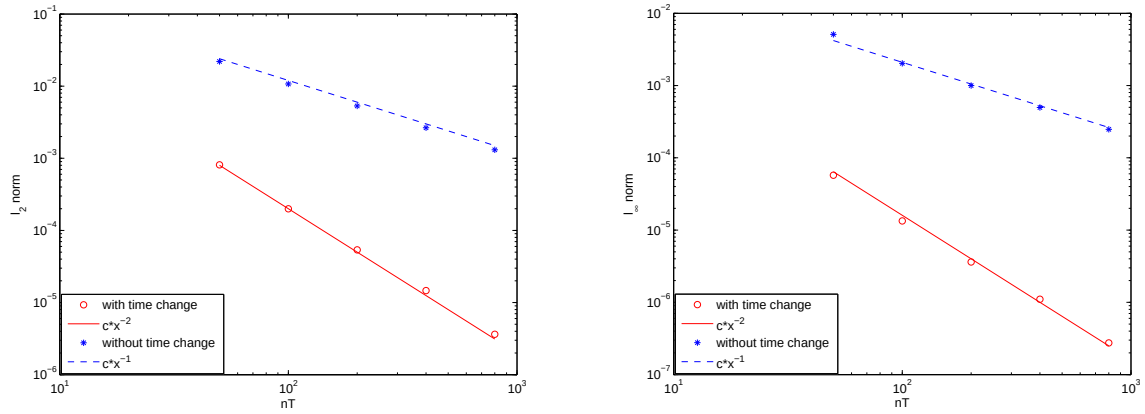
The convergence is clearly of first order without averaging and of second order with averaging.

Second, we modify the scheme using the idea from Reisinger and Whitley (2014) by changing the time variable $\tilde{t} = \sqrt{t}$. This change of variables leads to the new PDE

$$\frac{\partial V}{\partial \tilde{t}} + 2\tilde{t}\mathcal{L}V = 2\tau\chi(\tilde{t}^2, x),$$

instead of (2.58), to which we apply the numerical scheme.

In Figure 4.3, we show the convergence with and without time change, estimating the errors in a similar way to above, with $nX = 800$ fixed.



(a) l_2-norm. (b) l_∞-norm.

Figure 4.3: Convergence analysis for l_2- and l_∞-norms of the error depending on time-step with fixed mesh size.

The convergence is clearly of first order without time change and of second order with time change. We took here $T = 5$ to illustrate the effect more clearly.

4.3 Financial applications

4.3.1 Numerical experiments

In this section, we analyze the model characteristics and the impact of jumps. Specifically, we compute joint and marginal survival probabilities, CDS and FTD spreads as well as CVA and DVA depending on initial asset values. We also compute the difference between the solution with and without jumps.

Consider the parameters in Table 4.1.

$L_{1,0}$	$L_{2,0}$	$L_{12,0}$	$L_{21,0}$	R_1	R_2	T	σ_1	σ_2	ρ	ς_1	ς_2
60	70	10	15	0.4	0.45	1	1	1	0.5	1	1

Table 4.1: Model parameters.

For the model with jumps, we further consider the parameters in Table 4.2.

λ_1	λ_2	λ_{12}
0.5	0.5	0.3

Table 4.2: Jump intensities.

We compute all tests using a 100×100 spatial grid with the maximum values $X_1^{100} = X_2^{100} = 10$ and constant time step $\Delta\tau = 0.01$. As the parameters of the HV scheme, we choose $\sigma = \frac{1}{2}$ and $\theta = \frac{3}{4}$.

In Figures 4.4–4.6 we present various model characteristics and compare the results with and without jumps. From these figures, we can observe that jumps can have a significant impact, especially near the default boundaries:

- in Figure 4.4 for the joint survival probability, the biggest impact of jumps is around the default boundaries for both x_1 and x_2;
- in Figure 4.5 for the marginal survival probability of the first bank, we can observe that the biggest impact of jumps is near the default boundary of the first bank;
- for the CDS spread, in Figure 4.6, (b), the biggest impact of jumps is also seen near the default boundary, but it has the opposite direction, because jumps can only increase the CDS spread;
- in Figure 4.6, (d) for FTD the spread, the biggest impact of jumps is near both default boundaries, and it has a positive impact;
- finally, for CVA, (f), the highest impact of jumps is near the default boundary of the first bank, see Figure 4.6.

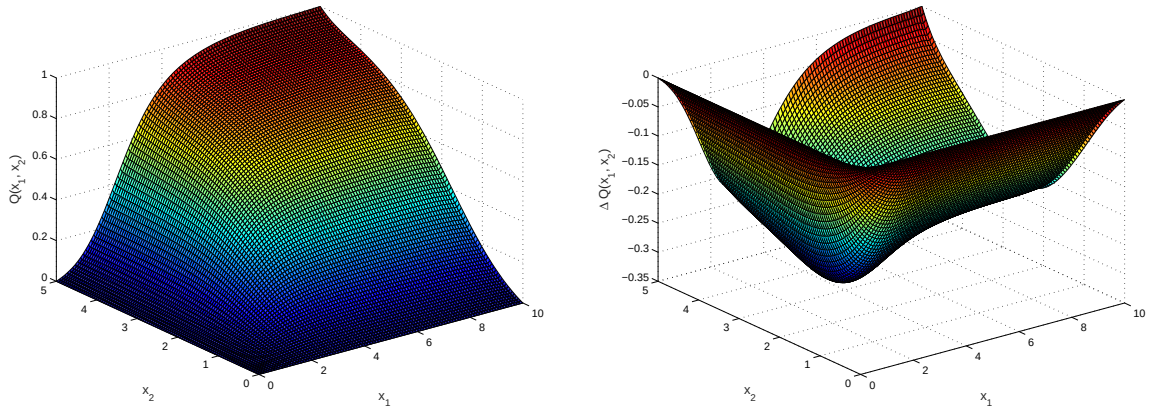
  The computation time for each model characteristic on the chosen grid is 1.5-1.7 s.[1]

4.3.2 Calibration

In this section we present calibration results of the model. There are eight unknown parameters, see (A.26)–(A.29): $\sigma_1, \sigma_2, \rho, \varsigma_1, \varsigma_2, \lambda_1, \lambda_2, \lambda_{12}$. We use CDS and equity put option prices (with different strikes) as market data. If FTD contracts are available, one can use them to estimate ρ and λ_{12}. Otherwise, historical estimation with share prices time series can be used.

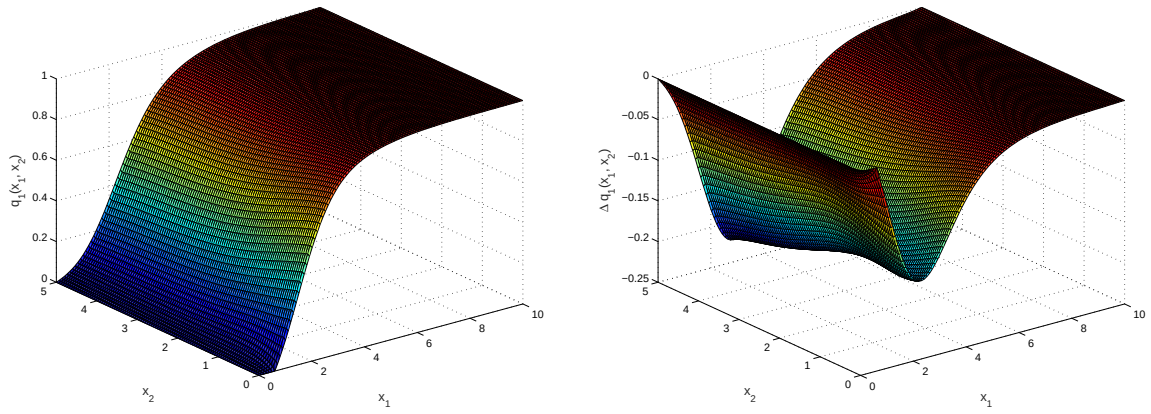
The data for external liabilities can be found in banks' balance sheets, which are publicly available. Usually, mutual liabilities data are not public information, thus we made an assumption that they are a fixed proportion of the total liabilities, which coincides

[1] We performed the scheme in Matlab on a PC with Intel Core i7 3.4Ghz CPU.



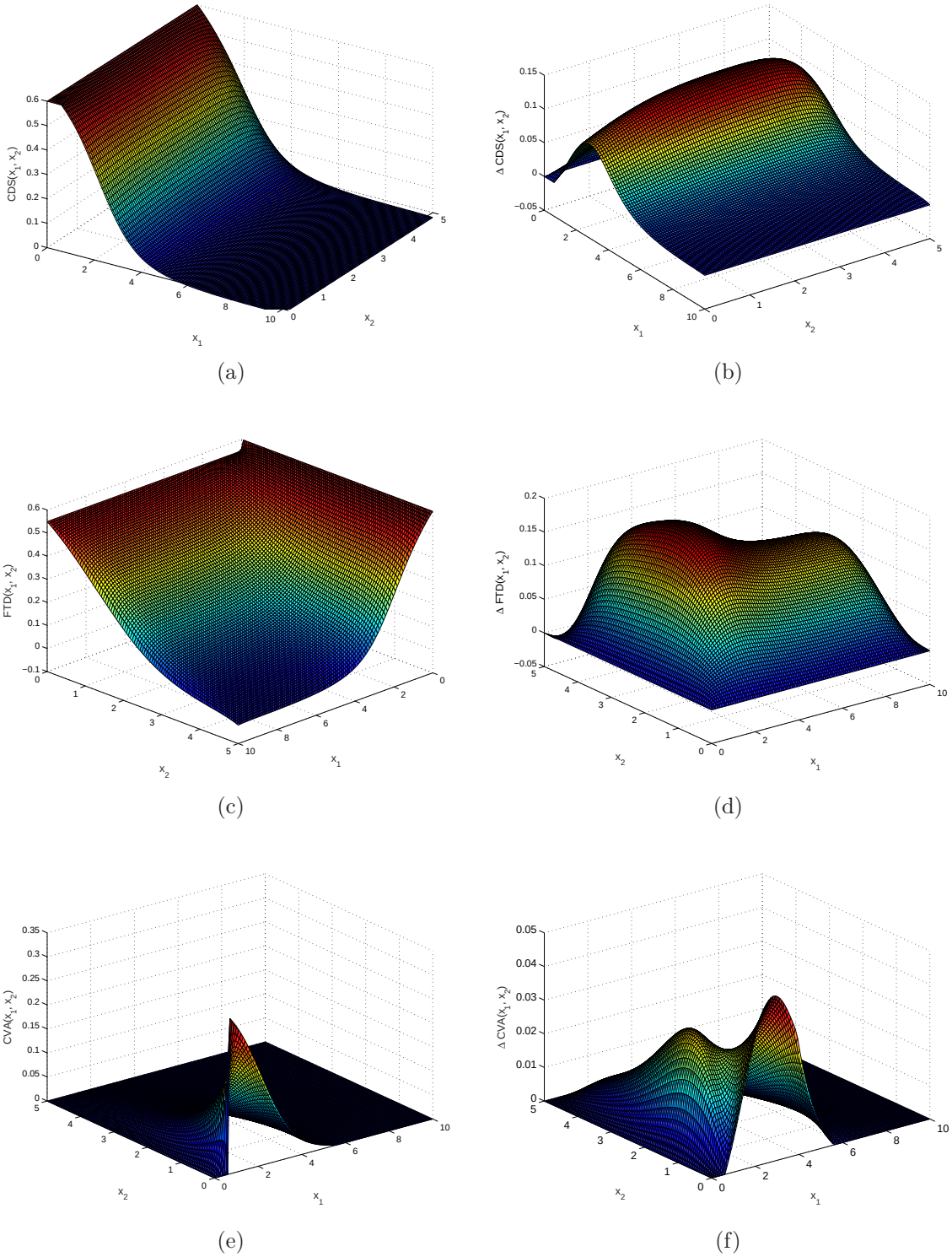
(a) (b)

Figure 4.4: The joint survival probability: (a) value, (b) difference between model with and without jumps.



(a) (b)

Figure 4.5: The marginal survival probability: (a) value, (b) difference between model with and without jumps.



(a) (b) (c) (d) (e) (f)

Figure 4.6: Values of different credit products with left the value and right the difference between model with and without jumps. Top row: Credit Default Swap spread, written on the first bank. Middle row: First-to-Default spread. Bottom row: CVA of CDS, where the first bank is Reference name (RN) and the second bank is Protection Seller (PS).

with David and Lehar (2014). In particular, we fix the mutual liabilities as 5% of total liabilities.

The asset's value is the sum of liabilities and equity price.

We choose Unicredit Bank as the first bank and Santander as the second bank. In Table 4.3 we provide their equity price E_i, assets A_i and liabilities L_i. As in Lipton and Sepp (2013), the liabilities are computed as a ratio of total liabilities and shares outstanding.

$E_1(0)$	$L_1(0)$	$A_1(0)$	$E_2(0)$	$L_2(0)$	$A_2(0)$
6.02	137.70	143.72	6.23	86.41	92.64

Table 4.3: Assets and liabilities on 30/06/2015 (Bloomberg).

For the calibration we choose 1-year at-the-money, in-the-money, and out-of-the-money equity put options on the banks, and 1-year CDS contracts. Since the spreads of CDS are usually significantly lower than the option prices, we scale them by some weight w_i in the objective function. As a result, we have the following 6-dimensional minimization problem:

$$\min_{\theta}\{w_1(V_1^{CDS}(\theta)-\bar{V}_1^{CDS})^2+\sum_{i=1}^{3}(V_1^{opt}(K_{i,1},\theta)-\bar{V}_1^{opt}(K_{i,1}))^2+ \\ +w_2(V_2^{CDS}(\theta)-\bar{V}_2^{CDS})^2+\sum_{i=1}^{3}(V_2^{opt}(K_{i,2},\theta)-\bar{V}_2^{opt}(K_{i,2}))^2\}, \quad (4.75)$$

where $\theta=(\sigma_1,\sigma_2,\lambda_1,\lambda_2,\varsigma_1,\varsigma_2)$, $V_i^{CDS}(\theta)$ is the model CDS spread on the i-th bank and $\bar{V}_i^{CDS}$ is the market CDS spread on the i-th bank, $V_1^{opt}(K,\theta)$ is the model price of the equity put option on the i-th bank with the strike K and $\bar{V}_i^{opt}(K)$ is the market price of the equity put option on the i-th bank with strike K. Strikes $K_{1,j},K_{2,j}$, and $K_{3,j}$ are chosen in such a way to take into account the smile. In particular, we choose $K_{1,j}=1.1E_j, K_{2,j}=E_j, K_{3,j}=0.9E_j$.

In order to find the global minimum of (4.75) by a Newton-type method, we need to find a good starting point, otherwise an optimization procedure might finish in local minima which are not global minima. To choose the starting point, we calibrate one-dimensional models for each bank without mutual liabilities

$$\min_{\theta_j}\{w_j(V_j^{CDS}(\theta_j)-\bar{V}_j^{CDS})^2+(V_j^{opt}(K_{1,j},\theta_j)-\bar{V}_j^{opt}(K_{1,j}))^2+ \\ +(V_j^{opt}(K_{2,j},\theta_j)-\bar{V}_i^{opt}(K_{2,j}))^2+(V_j^{opt}(K_{3,j},\theta_j)-\bar{V}_j^{opt}(K_{3,j}))^2\}, \quad (4.76)$$

where $\theta_j=(\sigma_j,\lambda_j,\varsigma_j)$ for $j=1,2$.

The global minima of (4.76) can be found via the **chebfun toolbox** (Driscoll et al. (2014)) that uses Chebyshev polynomials to approximate the function, and then the global minima can be easily found. The calibration results of the one-dimensional model for the first and the second banks are presented in Table 4.4. We note that the global

minima of (4.75) cannot be found via the chebfun toolbox, since it works with functions up to three variables. There are also more fundamental complexity issues for higher-dimensional tensor product interpolation.

σ_1	λ_1	ς_1	σ_2	λ_2	ς_2
0.0117	0.1001	0.3661	0.0154	0.0160	0.0545

Table 4.4: Calibrated parameters of one-dimensional models on 30/06/2015 for $T = 1$.

Similar to Lipton and Sepp (2013), for simplicity, we further assume that

$$\lambda_{\{12\}} = \rho \cdot \min(\lambda_1, \lambda_2). \tag{4.77}$$

Then, we estimate ρ from historical data. We take one year daily equity prices $E_i(t)$ by time series (from Bloomberg) and estimate the covariance of asset returns $r_t^i = \frac{\Delta A_i(t)}{A_i(t)}$

$$\widehat{\text{cov}}(A_1, A_2) = \sum_{i=1}^{n} \left(r_{i,1} - \bar{r_1}\right)\left(r_{i,2} - \bar{r}_2\right), \tag{4.78}$$

where $\bar{r}_1$ and $\bar{r}_2$ are the sample mean of asset returns.

Using (2.32), we can see that (4.78) converges to

$$\widehat{\text{cov}}(A_1, A_2) \underset{n\to+\infty}{\longrightarrow} \sigma_1\sigma_2\left(\rho + \lambda_{\{12\}}/(\varsigma_1\varsigma_2)\right). \tag{4.79}$$

Using the last equation and (4.77), we can extract the estimated values of ρ and $\lambda_{\{12\}}$. The estimation results are in Table 4.5.

	ρ	$\lambda_{\{12\}}$
Estimated value	0.510	0.0188
Confidence interval [2]	(0.500, 0.526)	(0.0182, 0.0194)

Table 4.5: Historically estimated correlation coefficients on 30/06/2015 with 1 year window.

Finally, we perform a six-dimensional (constrained) optimization of (4.75) with the starting point from Table 4.4 and correlation parameters from Table 4.5. We choose different alternatives of mutual liabilities to have a clear picture how mutual liabilities influence on model parameters. We use the **lsqnonlin** method in Matlab that uses a Trust Region Reflective algorithm Conn et al. (2000) (with the gradient computed numerically). Results are presented in Table 4.6.

Model	σ_1	λ_1	ς_1	σ_2	λ_2	ς_2
With jumps	0.0122	0.0950	0.3958	0.0160	0.0148	0.0505
Without jumps	0.0206	–	–	0.0317	–	–

Table 4.6: Calibrated parameters of two-dimensional model with mutual liabilities on 30/06/2015 for $T = 1$.

[2] We use a 3σ confidence interval.

In Table 4.7 we present joint and marginal survival probabilities computed using the equations from Section A.2. From these results, we can conclude that jumps play an important role in the model.

Model	Joint s/p	Marginal s/p
With jumps	0.9328	0.9666
Without jumps	0.9717	0.9801

Table 4.7: Joint and marginal survival probabilities for the calibrated models.

The computation time for our calibration procedure is 500s (75s for calibration of each one-dimensional model and 23s per iteration of the optimization procedure for the two-dimensional model that converges in 15 iterations).[3]

[3]We performed the calibration in Matlab on a PC with Intel Core i7 3.4Ghz (4 cores) CPU.

Chapter 5

Simulation of the infinite-dimensional model

5.1 Introduction

In this chapter, we study the infinite-dimensional approximation of the model, which we considered in Section 2.3. As we showed before, when the number of banks goes to infinity, the distance-to-default follows (2.80), a McKean–Vlasov equation with feedback from hitting the boundary.

We therefore propose and analyse numerical schemes for the simulation of the McKean–Vlasov equation

$$X_t = X_0 + W_t - \alpha L_t, \qquad t \in [0,T], \tag{5.1}$$

$$L_t = \mathbb{P}(\tau \le t), \qquad t \in [0,T], \tag{5.2}$$

$$\tau = \inf\{t \in [0,T] : X_t \le 0\}, \tag{5.3}$$

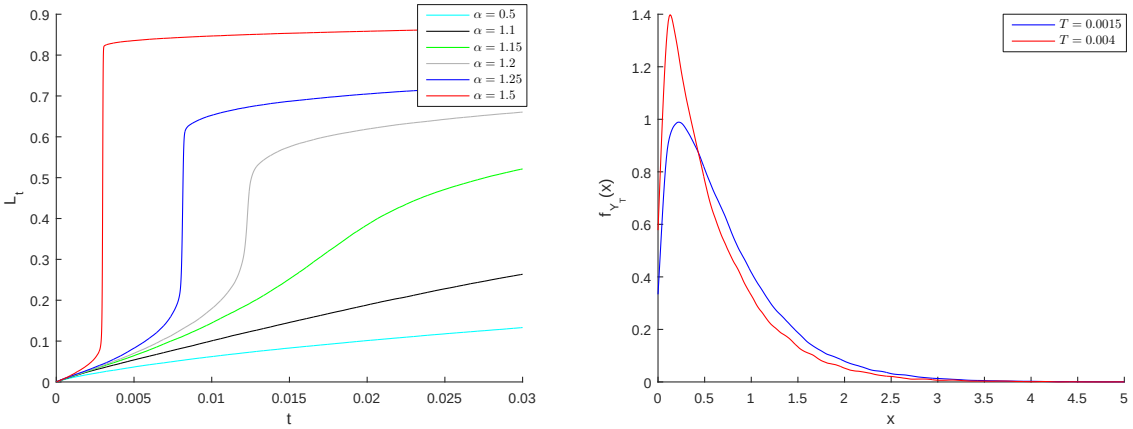
where $\alpha, T \in \mathbb{R}_+$, W a standard Brownian motion on a probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\ge 0}, \mathbb{P})$, on which is also given an $\mathbb{R}_+$-valued random variable X_0 independent of W. The non-linearity arises from the dependence of L_t in (5.1) on the law of X. More specifically, if $t \to L_t$ has a derivative p_τ, which is then the density of the hitting time τ of zero,

$$\mathrm{d}X_t = \mathrm{d}W_t - \alpha\, p_\tau(t)\,\mathrm{d}t, \qquad t \in (0,T),$$

so that the drift depends on the law of the path of X.

Theoretical properties of (5.1)–(5.3) have been studied in Hambly et al. (2018), who prove the existence of a differentiable solution $(L_t)_{0<t<t_*}$ up to an "explosion time" t_*. Conversely, they show that L cannot be continuous for all t for α above a threshold determined by the law of X_0. Such systemic events where discontinuities occur are also referred to as "blow-ups" in the literature.

The question of the constructive solution, however, remained open. Examples of a numerical solution computed with Algorithm 4 introduced in Section 5.2 are shown in Figure 5.1. The left plot shows the formation of a discontinuity in the loss function $t \to L_t$ for increasing α, with $X_0 \sim \mathrm{Gamma}(1.5, 0.5)$. The density of X_T for T before and after the shock is displayed in the right panel.



(a) (b)

Figure 5.1: (a) L_t for different α near the jump; (b) distribution of X_T for $X_T > 0$ before and after the jump (for $\alpha = 1.5$). Fitted by kernel density estimation with normal kernel for $N = 10^7$.

A similar model has been studied in Nadtochiy and Shkolnikov (2017), where the authors consider $\log(1 - L_t)$ instead of $-L_t$ in (5.1). Our numerical scheme can be applied in principle to this problem, but we concentrate the analysis on (5.1)–(5.3) in this chapter.

An earlier version of this problem is found in neuroscience, where a large network of electrically coupled neurons can be described by McKean–Vlasov type equations (Cáceres et al. (2011); Carrillo et al. (2013); Delarue et al. (2015b,a)). If a neuron's potential reaches some fixed threshold, it jumps to a higher potential level and sends a signal to other neurons. This feedback leads to the following equations

$$X_t = X_0 + \int_0^t b(X_s)\,\mathrm{d}s + \alpha\mathbb{E}[M_t] + W_t - M_t, \tag{5.4}$$

$$M_t = \sum_{k\geq 1} \mathbb{1}_{[0,t]}(\tau_k), \quad \tau_k = \inf\{t > \tau_{k-1} : X_{t-} \geq 1\}, \tag{5.5}$$

where $X_0 < 1$ a.s. The similarity to (5.1)–(5.3) is seen by noticing that $L_t = \mathbb{E}[\mathbb{1}_{[0,t]}(\tau)]$ and M_t in (5.5) is constant between hitting times, however, while in (5.4) an upper boundary is hit and after that the value resets to zero, in our model we are interested in hitting the zero boundary (from above), and after hitting the particle's value remains zero.

While McKean–Vlasov equations are an active area of current research, to our knowledge, there is only a fairly small number of papers where their simulation is studied rigorously and none of these encompass the models above.

Early works include Bossy and Talay (1997), which proves convergence (with order $1/2$ in the timestep and inverse number of particles) of a particle approximation to the distribution function for the measure ν_t of X_t in the classical McKean-Vlasov equation

$$X_t = X_0 + \int_{\mathbb{R}} \beta(X_t, u)\ \nu_t(\mathrm{d}u)\ \mathrm{d}t + \int_{\mathbb{R}} \alpha(X_t, u)\ \nu_t(\mathrm{d}u)\,\mathrm{d}W_t \tag{5.6}$$

with sufficient regularity. The proven rate in the timestep is improved to 1 in Antonelli et al. (2002) using Malliavin calculus techniques.

More recently, multilevel simulation algorithms have been proposed and analysed: Ricketson (2015) considers the special case of an SDE whose coefficients at time t depend on X_t and the expected value of a function of X_t; Szpruch et al. (2017) study a method based on fixed point iteration for the general case (5.6). An alternative variance reduction technique by importance sampling is given in Reis et al. (2018).

The system (5.1)–(5.3) above does not fall into the setting of (5.6) due to the extra path-dependence of the coefficients through the hitting time distribution. In this chapter, we therefore propose and analyse a particle scheme for (5.1)–(5.3) with an explicit timestepping scheme for the nonlinear term. We simulate N exchangeable particles at discrete time points with distance h, whereby at each time point t we use an estimator of L_{t-h} from the previous particle locations to approximate (5.3). We prove the convergence of the numerical scheme up to some time T as the time step h goes to zero and number of particles goes to infinity. We also prove that the scheme can be extended up to the explosion time under certain conditions on the model parameters. The order in h for this standard estimator is $1/2$. Next, we use Brownian bridges to better approximate the hitting probability, similar to barrier option pricing (Glasserman (2013)). In this case, the convergence order improves to $(1+\beta)/2$, where $\beta \in (0,1]$ is the Hölder exponent of the density of the initial value X_0 (e.g., 0.5 in the example from Figure 5.8 below). The order can be improved to 1 for all $\beta \in (0,1]$ by non-uniform time-stepping.

A main contribution is the first provably convergent scheme for equations of the type (5.1)–(5.3). The analysis uses a direct recursion of the error and regularity results proven in Hambly et al. (2018). This has the advantage that sharp convergence orders – i.e., consistent with the numerical tests – can be given, but also means that it seems difficult to apply the analysis directly to variations of the problem where such results are not available. Nonetheless, the method itself is natural and applicable in principle to other settings such as those outlined above.

The rest of the chapter is organized as follows. In Section 5.2 we list the running assumptions and state the main results; in Sections 5.3.1 and 5.3.2 we prove the uniform convergence of the discretized process; in Section 5.3.3, we show the convergence of Monte Carlo particle estimators with an increasing number of samples; in Section 5.3.4 we prove the convergence order for the scheme with Brownian bridges; in Section 5.4 we give numerical tests of the schemes; finally, in Section 5.5, we conclude.

5.2 Assumptions and main results

We begin by listing the assumptions. The first one, Hölder continuity at 0 of the initial density, is key for the regularity of the solution. The Hölder exponent will also limit the rate of convergence of the discrete time schemes.

Assumption 5.1. *We assume that X_0 has a density f_{X_0} supported on $\mathbb{R}_+$ such that*

$$f_{X_0}(x) \leq Bx^\beta, \qquad x \geq 0, \tag{5.7}$$

for some $\beta \in (0, 1]$.

Under Assumption 5.1, we can refer to Theorem 1.8 in Hambly et al. (2018) for the existence of a unique, differentiable solution $t \to L_t$ for (5.1)–(5.3) up to time

$$t_* := \sup \{t > 0 : ||L||_{H^1(0,t)} < \infty\} \in [0, \infty],$$

and a corresponding $\hat{B}$ such that for every $t < t_*$

$$L'_t \leq \hat{B} t^{-\frac{1-\beta}{2}} \ a.e. \tag{5.8}$$

This estimate admits a singularity of the rate of losses at time 0, however, what is actually observed in numerical studies (see Figure 5.1, left) is that the loss rate is bounded initially but then has a sharp peak for small β (and then especially for large α).

Integrating (5.8), we have for future reference a bound on L_t,

$$L_t \leq \tilde{B} t^{\frac{1+\beta}{2}}, \tag{5.9}$$

where $\tilde{B} = 2\hat{B}/(1 + \beta)$.

The following assumption will be used to control the propagation of the discretisation error, by bounding the density (especially at 0) of the running minimum of X and its approximations.

Assumption 5.2. *We assume that* $T < \min(T^*, t_*)$*, where* T^* *is defined by*

$$\alpha B \left[\sqrt{\frac{2T^*}{\pi}} + \alpha \tilde{B} (T^*)^{\frac{1+\beta}{2}} \right]^{\beta} = 1, \tag{5.10}$$

with B *and* $\tilde{B}$ *the smallest constants such that* (5.7) *and* (5.9) *hold for given* β.

In the following, we assume that Assumptions 5.1 and 5.2 hold. Consider a uniform time mesh $0 = t_0 < t_1 < \ldots < t_n = T$, where $t_i - t_{i-1} = h$, and a discretized process, for $1 \leq i \leq n$,

$$\begin{aligned} \tilde{X}_{t_i} &= X_0 + W_{t_i} - \alpha \tilde{L}_{t_i}, & (5.11) \\ \tilde{L}_{t_i} &= \mathbb{P}(\tilde{\tau} < t_i), & (5.12) \\ \tilde{\tau} &= \min_{0 \leq j \leq n} \{\tilde{X}_{t_j} \leq 0\}. & (5.13) \end{aligned}$$

We extend $\tilde{L}_{t_i}$ to $[0, T]$ by setting $\tilde{L}_s = \tilde{L}_{t_{i-1}}$ for $t_{i-1} < s < t_i$.

The first theorem, proven in Section 5.3.1, shows that $\tilde{L}_t$ converges uniformly to L_t.

Theorem 5.1. *Consider* $\tilde{L}_{t_i}$ *from* (5.11)–(5.13) *and* L_t *from* (5.1)–(5.3). *Then, for any* $\delta > 0$*, there exists* $C > 0$ *independent of* h *such that*

$$\max_{i \leq n} |\tilde{L}_{t_i} - L_{t_i}| \leq C h^{\frac{1}{2}-\delta}. \tag{5.14}$$

We now propose a particle simulation scheme for (5.11)–(5.13) in Algorithm 4.

Algorithm 4 Discrete time Monte Carlo scheme for simulation of the loss process

Require: N — number of Monte Carlo paths
Require: n — number of time steps: $0 < t_1 < t_2 < \ldots < t_n$
1: Draw N samples of X_0 (from initial distribution) and W (a Brownian path)
2: Define $\hat{L}_0 = 0$
3: **for** $i = 1 : n$ **do**
4: Estimate $\tilde{L}_{t_i}$ by $\hat{L}^N_{t_i} = \frac{1}{N}\sum_{k=1}^{N} \mathbb{1}_{\{\min_{j<i} \hat{X}^{(k)}_{t_j} \leq 0\}}$
5: **for** $k = 1 : N$ **do**
6: Update $\hat{X}^{(k)}_{t_i} = X^{(k)}_0 + W^{(k)}_{t_i} - \alpha \hat{L}^N_{t_i}$
7: **end for**
8: **end for**

In Section 5.3.3, we prove convergence in probability of Algorithm 4 as $N \to \infty$.

Theorem 5.2. *For all $i \leq n$,*

$$\hat{L}_{t_i} \xrightarrow[N\to\infty]{\mathbb{P}} \tilde{L}_{t_i}. \tag{5.15}$$

Next, we improve our scheme by using a Brownian bridge strategy to estimate the hitting probabilities. In order to do this, we consider the process

$$\breve{X}_t = X_0 + W_t - \alpha \breve{L}_t, \qquad t \in [t_i, t_{i+1}), \tag{5.16}$$
$$\breve{L}_t = \mathbb{P}(\breve{\tau} < t_i), \qquad t \in [t_i, t_{i+1}), \tag{5.17}$$
$$\breve{\tau} = \inf_{0\leq s\leq T}\{\breve{X}_s \leq 0\}. \tag{5.18}$$

Then, for each Brownian path $(W^{(k)}_t)_{t\geq 0}$, we compute $\bar{X}^{(k)}_t = X^{(k)}_0 + W^{(k)}_t - \alpha \bar{L}^N_t$ in (t_i, t_{i+1}), where $\bar{L}^N_t$ is an N-sample estimator of $\breve{L}_{t_i}$ given below. Then, we estimate the hitting probability

$$\begin{aligned} p^{(k)}_{t_i} &= \mathbb{P}\left(\inf_{s<t_i} \bar{X}^{(k)}_s > 0 | \bar{X}^{(k)}_0, \ldots, \bar{X}^{(k)}_{t_i}\right) \\ &= \prod_{j=1}^{i} \mathbb{P}\left(\inf_{s\in[t_{j-1},t_j)} \bar{X}^{(k)}_s > 0 \mid \bar{X}^{(k)}_{t_{j-1}}, \bar{X}^{(k)}_{t_j}\right). \end{aligned}$$

It is well-known that using Brownian bridge, one can derive the hitting probability of Brownian motion:

$$\mathbb{P}(\inf_{s\in[0,t]} W_s \leq 0 \mid W_0 > 0, W_t > 0) = \exp\left(-\frac{2W_t W_0}{T}\right).$$

Using the last result, we have

$$p^{(k)}_{t_i} = \prod_{j=1}^{i}\left(1 - \exp\left(-\frac{2(\bar{X}^{(k)}_{t_{j-1}} \vee 0)(\bar{X}^{(k)}_{t_j-} \vee 0)}{h}\right)\right).$$

Thus, a natural choice for $\bar{L}^N_{t_i}$ is

$$\bar{L}^N_{t_i} = \frac{1}{N}\sum_{k=1}^{N}\left(1 - p^{(k)}_{t_i}\right). \tag{5.19}$$

As a result, the new algorithm with the Brownian bridge modification is the following.

Algorithm 5 Discrete time Monte Carlo scheme with Brownian bridge

Require: N — number of Monte Carlo paths
Require: n — number of time steps: $0 < t_1 < t_2 < \ldots < t_n$
1: Draw N samples X_0 (from the initial distribution) and W (a Brownian path)
2: **for** $i = 1 : n$ **do**
3: Estimate $\bar{L}_{t_i}$ using (5.19)
4: **for** $k = 1 : N$ **do**
5: Update $\bar{X}^{(k)}_{t_i} = X^{(k)}_0 + W^{(k)}_{t_i} - \alpha \bar{L}^N_{t_i}$
6: **end for**
7: **end for**

The convergence rate for (5.17) is given as follows and proven in Section 5.3.4.

Theorem 5.3. *Consider $\breve{L}_t$ from (5.16)–(5.18) and L_t from (5.1)–(5.3), $\beta \in (0, 1]$ from Assumption 5.1. Then, there exists $C > 0$ independent of h such that*

$$\max_{i \le n} |\breve{L}_{t_i} - L_{t_i}| \le C h^{\frac{1+\beta}{2}}. \tag{5.20}$$

We will later give a result with variable time steps which achieves rate 1 for all β.

5.3 Convergence results

5.3.1 Convergence of the timestepping scheme

In this section we prove Theorem 5.1. Then, in Section 5.3.2, under a modification of Assumption 5.2, we formulate and prove an improvement of this theorem which extends the applicable time interval.

The proof is based on induction on the error bound over the timesteps, which requires an error estimate of the hitting probability after discretisation (Lemmas 5.1 and 5.2), a sort of consistency, plus a control of the resulting misspecification of the barrier through a bound on the density of the running minimum (Lemma 5.3), a kind of stability.

As we have crude estimates on the densities, using only Assumption 5.1 but no sharper bound on, and regularity of, f_{X_0}, or any regularity of the distribution of $\inf_{s \le t}(W_s - \alpha L_s)$ or its approximations, the time T^* until which the numerical scheme is shown to converge will by no means be sharp. Indeed, in our numerical tests (see Section 5.4) we did not encounter any difficulties for any t, even $t > t_*$.

We first formulate and prove these auxiliary results which we will use to prove Theorems 5.1, 5.3, and 5.4.

First, we modify Proposition 1 from Asmussen et al. (1995), where an analogous result is shown for standard Brownian motion, without the presence of L and $\tilde{L}$ (i.e., $\alpha = 0$).

Lemma 5.1. *Define* $X_t^* = X_0 + W_{h\lfloor \frac{t}{h} \rfloor} - \alpha L_{h\lfloor \frac{t}{h} \rfloor}$ *and* $\tilde{X}_t^* = X_0 + W_{h\lfloor \frac{t}{h} \rfloor} - \alpha \tilde{L}_{h\lfloor \frac{t}{h} \rfloor}$ *for* $t < t_*$. *Then, as* $h \to 0$,

$$\frac{1}{\sqrt{h}} \sup_{s \le t} (X_s - X_s^*) \to_d \sqrt{2 \log \frac{t}{h}}, \tag{5.21}$$

and

$$\frac{1}{\sqrt{h}} \sup_{s \le t} \left(\tilde{X}_s - \tilde{X}_s^* \right) \to_d \sqrt{2 \log \frac{t}{h}}. \tag{5.22}$$

Proof. Writing

$$\begin{aligned} \sup_{s \le h\lfloor \frac{t}{h} \rfloor} (X_s - X_s^*) &= \max_{1 \le k \le \lfloor \frac{t}{h} \rfloor} \sup_{0 \le s \le h} \left(W_{h(k-1)+s} - W_{h(k-1)} - \alpha (L_{h(k-1)+s} - L_{h(k-1)}) \right) \\ &= \max_{1 \le k \le \lfloor \frac{t}{h} \rfloor} \sup_{0 \le s \le h} \left(W_{h(k-1)+s} - W_{h(k-1)} - \alpha s L'_{h(k-1)+\theta s} \right) \end{aligned}$$

for some $\theta \in [0,1]$, the last expression can be estimated from both sides,

$$\begin{aligned} &\max_{1 \le k \le \lfloor \frac{t}{h} \rfloor} \left\{ \sup_{0 \le s \le h} \left(W_{h(k-1)+s} - W_{h(k-1)} \right) - \alpha \sup_{0 \le s \le h} h L'_{h(k-1)+\theta s} \right\} \le \\ &\qquad \max_{1 \le k \le \lfloor \frac{t}{h} \rfloor} \sup_{0 \le s \le h} \left(W_{h(k-1)} - W_{h(k-1)+s} - \alpha s L'_{h(k-1)+\theta s} \right) \\ &\qquad \le \max_{1 \le k \le \lfloor \frac{t}{h} \rfloor} \left\{ \sup_{0 \le s \le h} \left(W_{h(k-1)} - W_{h(k-1)+s} \right) - \alpha \inf_{0 \le s \le h} s L'_{h(k-1)+\theta s} \right\}. \end{aligned}$$

Since $0 \le L'_s \le \hat{B} s^{-\frac{1-\beta}{2}}$,

$$\begin{aligned} \sup_{s \le h\lfloor \frac{t}{h} \rfloor} (X_s - X_s^*) &\ge \max_{1 \le k \le \lfloor \frac{t}{h} \rfloor} \left\{ \sup_{0 \le s \le h} \left(W_{h(k-1)+s} - W_{h(k-1)} \right) - \alpha \hat{B} h (h(k-1))^{-\frac{1-\beta}{2}} \right\} \\ &=_d \sqrt{h} \max_{1 \le k \le \lfloor \frac{t}{h} \rfloor} \left\{ \sup_{0 \le s \le 1} W_s^{(k)} - \alpha \hat{B} \sqrt{h} (h(k-1))^{-\frac{1-\beta}{2}} \right\}, \end{aligned}$$

and

$$\begin{aligned} \sup_{s \le h\lfloor \frac{t}{h} \rfloor} (X_s - X_s^*) &\le \max_{1 \le k \le \lfloor \frac{t}{h} \rfloor} \left\{ \sup_{0 \le s \le h} \left(W_{h(k-1)+s} - W_{h(k-1)} \right) \right\} \\ &=_d \sqrt{h} \max_{1 \le k \le \lfloor \frac{t}{h} \rfloor} \left\{ \sup_{0 \le s \le 1} W_s^{(k)} \right\}, \end{aligned}$$

where $W^{(1)}, W^{(2)}, \ldots$ are i.i.d. copies of W.

Taking into account that $\sqrt{h}(h(k-1))^{-\frac{1-\beta}{2}} \to 0$, as $h \to 0$, and applying the same arguments as in Proposition 1 in Asmussen et al. (1995), we get (5.21).

Similar,

$$\begin{aligned} \sup_{s \le h\lfloor \frac{t}{h} \rfloor} \left(\tilde{X}_s - \tilde{X}_s^* \right) &= \max_{1 \le k \le \lfloor \frac{t}{h} \rfloor} \sup_{0 \le s \le h} \left(W_{h(k-1)+s} - W_{h(k-1)} \right) \\ &= \max_{1 \le k \le \lfloor \frac{t}{h} \rfloor} \sup_{0 \le s \le h} \left(W_{h(k-1)+s} - W_{h(k-1)} \right) =_d \sqrt{h} \max_{1 \le k \le \lfloor \frac{t}{h} \rfloor} \left\{ \sup_{0 \le s \le 1} W_s^{(k)} \right\}, \end{aligned}$$

where $W^{(1)}, W^{(2)}, \ldots$ are i.i.d. copies of W.

Again, applying the same arguments as in Proposition 1 in Asmussen et al. (1995), we get (5.22). □

The next lemma deduces the convergence rate of the hitting probabilities on the mesh.

Lemma 5.2. *Consider the processes X and $\tilde{X}$. Then, for any $\delta > 0$ there exist $\gamma, \tilde{\gamma} > 0$, independent of h and i, such that*

$$0 \leq \mathbb{P}\left(\min_{j<i} X_{t_j} > 0\right) - \mathbb{P}\left(\inf_{s<t_i} X_s > 0\right) \leq \gamma h^{\frac{1}{2}-\delta} \tag{5.23}$$

and

$$0 \leq \mathbb{P}\left(\min_{j<i} \tilde{X}_{t_j} > 0\right) - \mathbb{P}\left(\inf_{s<t_i} \tilde{X}_s > 0\right) \leq \tilde{\gamma} h^{\frac{1}{2}-\delta}, \tag{5.24}$$

where $t_i < t_*$.

Proof. We first prove (5.23). It is obvious that

$$\mathbb{P}\left(\min_{j<i} X_{t_j} > 0\right) \geq \mathbb{P}\left(\inf_{s<t_i} X_s > 0\right).$$

Also,

$$\mathbb{P}\left(\min_{j<i} X_{t_j} > 0\right) = \mathbb{P}\left(\inf_{s<t_i} X_s > 0\right) + \mathbb{P}\left(\left\{\inf_{s<t_i} X_s \leq 0\right\} \cap \left\{\min_{j<i} X_{t_j} > 0\right\}\right).$$

Thus,

$$\begin{aligned}\mathbb{P}\left(\min_{j<i} X_{t_j} > 0\right) - \mathbb{P}\left(\inf_{s<t_i} X_s > 0\right) &= \mathbb{P}\left(\left\{\inf_{s<t_i} X_s \leq 0\right\} \cap \left\{\min_{j<i} X_{t_j} > 0\right\}\right) \\ &\leq \mathbb{P}\left(\min_{j<i} X_{t_j} - \inf_{s<t_i} X_s \geq \varepsilon\right) + \mathbb{P}\left(\min_{j<i} X_{t_j} \in (0,\varepsilon), \min_{j<i} X_{t_j} - \inf_{s<t_i} X_s < \varepsilon\right),\end{aligned}$$

for any $\varepsilon > 0$.

Consider the first term. Using Markov's inequality

$$\mathbb{P}\left(\min_{j<i} X_{t_j} - \inf_{s<t_i} X_s \geq \varepsilon\right) \leq \frac{\mathbb{E}\left[\left(\min_{j<i} X_{t_j} - \inf_{s<t_i} X_s\right)^p\right]}{\varepsilon^p},$$

for any $p \geq 1$.

Using Lemma 5.1,

$$\frac{1}{\sqrt{h}}\left(\min_{j<i} X_{t_j} - \inf_{s<t_i} X_s\right) \to_d \sqrt{2\log\frac{t_i}{h}}$$

thus

$$\frac{\mathbb{E}\left[\left(\min_{j<i} X_{t_j} - \inf_{s<t_i} X_s\right)^p\right]}{\varepsilon^p \cdot h^{p/2}} \longrightarrow \frac{(2\log\frac{t_i}{h})^{p/2}}{\varepsilon^p}.$$

For the second term

$$\mathbb{P}\left(\min_{j<i} X_{t_j} \in (0,\varepsilon), \min_{j<i} X_{t_j} - \inf_{s<t_i} X_s < \varepsilon\right) \leq \mathbb{P}\left(\min_{j<i} X_{t_j} \in (0,\varepsilon)\right) =$$
$$\mathbb{P}\left(\min_{j<i} X_{t_j} > 0\right) - \mathbb{P}\left(\min_{j<i} X_{t_j} > \varepsilon\right) = \bar{F}_i(\varepsilon) - \bar{F}_i(0) \leq \varepsilon \sup_{\theta\in[0,1]} \bar{\varphi}_i(\theta\varepsilon),$$

where $\bar{F}_i(x)$ and $\bar{\varphi}_i(x)$ are the CDF and PDF of the process $\min_{j<i} X_{t_j}$. We also note that $\bar{\varphi}_i(\theta\varepsilon)$ is bounded according to Lemma 5.3.

Combining both terms, we have

$$\mathbb{P}\left(\min_{j<i} X_{t_j} > 0\right) - \mathbb{P}\left(\inf_{s<t_i} X_s > 0\right) \leq A\frac{h^{p/2}}{\varepsilon^p} + B\varepsilon.$$

Minimising the right-hand side over ε, we find $\varepsilon = h^a$ with $a = \frac{p}{2(p+1)}$. As a result, choosing $p \to \infty$, we get that for any $\delta > 0$, there exists $\gamma > 0$, such that

$$\mathbb{P}\left(\min_{j<i} X_{t_j} > 0\right) - \mathbb{P}\left(\inf_{s<t_i} X_s > 0\right) \leq \gamma h^{\frac{1}{2}-\delta}.$$

For (5.24), we use identical steps and the corresponding estimates in Lemma 5.1 and Lemma 5.3 for $\tilde{X}$ instead of X.

□

Finally, we bound the probability that the running minimum is close to the boundary.

Lemma 5.3. *Consider $Z_i = \inf_{s\leq t_i} X_s$, $\bar{Z}_i = \min_{j<i} X_{t_j}$, and $\tilde{Z}_i = \min_{j<i} \tilde{X}_{t_j}$ for some $i \leq n$ and $t_i < t_*$. Then, $Z_i, \bar{Z}_i$, and $\tilde{Z}_i$ each have a density, denoted φ_i, $\bar{\varphi}_i$, and $\tilde{\varphi}_i$, respectively, with*

$$\max(\varphi_i(z), \bar{\varphi}_i(z), \tilde{\varphi}_i(z)) \leq B\left[(z\vee 0) + \sqrt{\frac{2t_i}{\pi}} + \alpha\tilde{B}t_i^{\frac{1+\beta}{2}}\right]^\beta, \tag{5.25}$$

where β, B are from (5.7) *and $\tilde{B}$ is from* (5.9).

Proof. We can rewrite $Z_i = X_0 + V_i$, where $V_i = \inf_{s\leq t_i}(W_s - \alpha L_s)$. As X_0 and V_i are independent, using convolution, we have

$$\varphi_i(z) = \int_{-\infty}^{+\infty} f_{X_0}(z-v)\mathbb{P}(V_i \in dv).$$

Using the fact that $V_i \leq 0, X_0 \geq 0$, and (5.7), we have

$$\varphi_i(z) = \int_{-\infty}^{z\wedge 0} f_{X_0}(z-v)\mathbb{P}(V_i \in dv) \leq B\int_{-\infty}^{z\wedge 0}(z-v)^\beta \mathbb{P}(V_i \in dv)$$
$$= BF_{V_i}(z\wedge 0)\int_{-\infty}^{z\wedge 0}(z-v)^\beta \frac{\mathbb{P}(V_i \in dv)}{F_{V_i}(z\wedge 0)}, \tag{5.26}$$

where F_{V_i} is the CDF of V_i. It is obvious that $F_{V_i}(z) > 0$ for all $z \leq 0$. Indeed, $F_{V_i}(z) \geq \mathbb{P}(V_i \leq z) \geq \mathbb{P}(\inf_{s\leq t_i} W_s \leq z) > 0$.

Since $\beta \in (0,1]$, for all z the function $(z-\cdot)^\beta : (-\infty, z\wedge 0) \to (0,\infty)$ is concave and, using Jensen's inequality for the proper probability measure $\frac{\mathbb{P}(V_i\in dv)}{F_{V_i}(z\wedge 0)}$ on $(-\infty, z\wedge 0)$,

$$\varphi_i(z) \le BF_{V_i}(z\wedge 0)^{1-\beta}\left(\int_{-\infty}^{z\wedge 0}(z-v)\mathbb{P}(V_i\in dv)\right)^\beta$$
$$= BF_{V_i}(z\wedge 0)^{1-\beta}\left(\mathbb{E}\left[(z-V_i)\mathbb{1}_{\{V_i\le z\}}\right]\right)^\beta, \quad (5.27)$$

where $(V_i \le z) \Leftrightarrow (V_i \le z\wedge 0)$ as $V_i \le 0$. Since $t \to L_t$ is non-decreasing, inserting V_i,

$$\varphi_i(z) \le BF_{V_i}(z\wedge 0)^{1-\beta}(\mathbb{E}[(z-\inf_{s\le t_i}\{W_s\}+\alpha L_{t_i})\mathbb{1}_{\{\inf_{s\le t_i}\{W_s\}-\alpha L_{t_i}\le z\}}])^\beta$$
$$\le BF_{V_i}(z\wedge 0)^{1-\beta}\left(\mathbb{E}[(z-\inf_{s\le t_i}\{W_s\}+\alpha\tilde{B}t_i^{\frac{1+\beta}{2}})\mathbb{1}_{\left\{\inf_{s\le t_i}\{W_s\}\le z+\alpha\tilde{B}t_i^{\frac{1+\beta}{2}}\right\}}]\right)^\beta. \quad (5.28)$$

By simple integration of the density of the running minimum of Brownian motion,

$$\mathbb{E}[(\xi-\inf_{s\le t_i}W_s)\mathbb{1}_{\{\inf_{s\le t_i}W_s\le\xi\}}] = \left(2\Phi\left(\frac{\xi}{\sqrt{t_i}}\right)\wedge 1\right)\xi + \sqrt{\frac{2t_i}{\pi}}\mathrm{e}^{-\frac{\xi^2}{2t_i}}, \quad (5.29)$$

and, taking $\xi = z+\alpha\tilde{B}t_i^{\frac{1+\beta}{2}}$, we get from (5.28) for $z>0$,

$$\varphi_i(z) \le B\left[z+\alpha\tilde{B}t_i^{\frac{1+\beta}{2}}+\sqrt{\frac{2t_i}{\pi}}\right]^\beta,$$

and, similarly, for $z\le 0$,

$$\varphi_i(z) \le B\left[\alpha\tilde{B}t_i^{\frac{1+\beta}{2}}+\sqrt{\frac{2t_i}{\pi}}\right]^\beta.$$

Combing the last two equations, we finally get (5.25) for φ_i.

Using similar arguments for $\bar{Z}_i$ and $\tilde{Z}_i$, using $\inf_{s\le t_i}W_s \le \min_{j<i}W_{t_j}$ and $L_{t_i}\ge\tilde{L}_{t_i}$, respectively, we get the analogous estimates for $\bar{\varphi}_i$ and $\tilde{\varphi}_i$, and hence (5.25). □

We are now in a position to prove Theorem 5.1.

Proof of Theorem 5.1. We shall proceed by induction. For $t_0=0$, we have $L_0=\tilde{L}_0$; for t_1, we have $\min_{j<1}\tilde{X}_{t_j}=X_0$ and $\mathbb{P}(X_0>0)=1$, hence,

$$|\tilde{L}_{t_1}-L_{t_1}| \le \left|1-\mathbb{P}\left(\inf_{s<t_1}X_s>0\right)\right|,$$

which can be estimated using Lemma 5.2, (5.23).

For $i>1$, assume now that we have shown $\tilde{L}_{t_j}=L_{t_j}-\tilde{C}_j h^{\frac{1}{2}-\delta}$ for $j<i$, where $\tilde{C}_j\ge 0$ as $\tilde{L}_{t_j}\le L_{t_j}$. We split the error into two contributions,

$$|\tilde{L}_{t_i}-L_{t_i}| = \left|\mathbb{P}\left(\min_{j<i}\tilde{X}_{t_j}>0\right)-\mathbb{P}\left(\inf_{s<t_i}X_s>0\right)\right|$$
$$\le \left|\mathbb{P}\left(\min_{j<i}\tilde{X}_{t_j}>0\right)-\mathbb{P}\left(\min_{j<i}X_{t_j}>0\right)\right| + \left|\mathbb{P}\left(\min_{j<i}X_{t_j}>0\right)-\mathbb{P}\left(\inf_{s<t_i}X_s>0\right)\right|.$$

We can use Lemma 5.2, (5.23), for the second term. For the first term,

$$\begin{aligned}
&\mathbb{P}\left(\min_{j<i}\tilde{X}_{t_j}>0\right)-\mathbb{P}\left(\min_{j<i}X_{t_j}>0\right)\\
&\qquad=\mathbb{P}\left(\min_{j<i}\left(X_{t_j}+\alpha\tilde{C}_j h^{\frac{1}{2}-\delta}\right)>0\right)-\mathbb{P}\left(\min_{j<i}X_{t_j}>0\right)\\
&\qquad\leq\mathbb{P}\left(\min_{j<i}X_{t_j}>-\alpha\max_{j<i}\tilde{C}_j h^{\frac{1}{2}-\delta}\right)-\mathbb{P}\left(\min_{j<i}X_{t_j}>0\right)\\
&\qquad=\bar{F}_i(0)-\bar{F}_i\left(-\alpha\max_{j<i}\tilde{C}_j h^{\frac{1}{2}-\delta}\right)\\
&\qquad\leq\alpha\max_{j<i}\tilde{C}_j h^{\frac{1}{2}-\delta}\sup_{\theta\in[0,1]}\bar{\varphi}_i\left(-\theta\alpha\max_{j<i}\tilde{C}_j h^{\frac{1}{2}-\delta}\right),
\end{aligned}$$

where $\bar{F}_i(x)$ and $\bar{\varphi}_i(x)$ are the CDF and pdf of $\min_{j<i}X_{t_j}$.

Then, using Lemma 5.3, we have

$$\bar{\varphi}_i\left(-\theta\frac{\alpha}{2}\max_{j<i}\tilde{C}_j h^{\frac{1}{2}-\delta}\right)\leq B\left[\sqrt{\frac{2t_i}{\pi}}+\alpha\tilde{B}t_i^{\frac{1+\beta}{2}}\right]^{\beta},$$

as $-\theta\frac{\alpha}{2}\max_{j<i}\tilde{C}_j h^{\frac{1}{2}-\delta}<0$. As a result, we have the following inequality for $\tilde{C}_i$,

$$\tilde{C}_i\leq\alpha\max_{j<i}\tilde{C}_j B\left[\sqrt{\frac{2t_i}{\pi}}+\alpha\tilde{B}t_i^{\frac{1+\beta}{2}}\right]^{\beta}+\gamma,$$

hence $\tilde{C}_i$ is bounded independent of i and h by Assumption 5.2. By induction we get (5.14).

□

5.3.2 Extension of the result in time

The following result extends the applicability of Theorem 5.1 up to the explosion time t_* under certain conditions on the parameters, as specified precisely in (5.31) below.

We shall adapt Theorem 1 in Borovkov and Novikov (2005), which states: For a Lipschitz function f with Lipschitz constant K, and g such that $\sup_{s\leq t}|f(s)-g(s)|\leq\varepsilon$,

$$|\mathbb{P}(\exists s\in[0,t]:W_s<f(s))-\mathbb{P}(\exists s\in[0,t]:W_s<g(s))|\leq\left(2.5K+\frac{2}{\sqrt{t}}\right)\varepsilon.$$

By Remark 2 in Borovkov and Novikov (2005), this result can be improved for a non-decreasing function g. Indeed, retracing the steps in their proof and using monotonicity, one finds easily the slightly better bound

$$|\mathbb{P}(\exists s\in[0,t]:W_s<f(s))-\mathbb{P}(\exists s\in[0,t]:W_s<g(s))|\leq 2\left(K+\frac{1}{\sqrt{t}}\right)\varepsilon.$$

In our case, we cannot directly apply the result with $f(s)=-X_0+\alpha L_s$ and $g(s)=-X_0+\alpha\tilde{L}_s$ as f is not guaranteed to be Lipschitz at $s=0$. But, along the lines of the proof of Theorem 1 in Borovkov and Novikov (2005), the above result can be modified as follows:

Lemma 5.4. *For a non-decreasing function g which is Lipschitz with constant K on $[T^*, T]$, and a function f such that $f(t) = g(t)$ for $t \le T^*$ and $\sup_{s\le T} |f(s) - g(s)| \le \varepsilon$,*

$$|\mathbb{P}(\exists s \in [0,T] : W_s < f(s)) - \mathbb{P}(\exists s \in [0,T] : W_s < g(s))| \le 2\left(K + \frac{1}{\sqrt{T}}\right)\varepsilon. \tag{5.30}$$

Theorem 5.4. *Assume T^* from (5.10) satisfies*

$$2\alpha\left(\hat{B}(T^*)^{-\frac{1-\beta}{2}} + \frac{1}{\sqrt{T^*}}\right) < 1. \tag{5.31}$$

Then Theorem 5.1 holds for any $T < t_*$ (and not only up to T^* as per Assumption 5.2).

Proof. We first split again the error by

$$\begin{aligned}|\tilde{L}_{t_i} - L_{t_i}| &= \left|\mathbb{P}\left(\min_{j<i} \tilde{X}_{t_j} > 0\right) - \mathbb{P}\left(\inf_{s\le t_i} X_s > 0\right)\right| \\ &\le \left|\mathbb{P}\left(\inf_{s\le t_i} \tilde{X}_s > 0\right) - \mathbb{P}\left(\inf_{s\le t_i} X_s > 0\right)\right| + \left|\mathbb{P}\left(\min_{j<i} \tilde{X}_{t_j} > 0\right) - \mathbb{P}\left(\inf_{s\le t_i} \tilde{X}_s > 0\right)\right|.\end{aligned} \tag{5.32}$$

The second term can be estimated from Lemma 5.2, (5.24). Again, we shall then proceed by induction. We have already shown that $|\tilde{L}_{t_i} - L_{t_i}| \le C_i h^{\frac{1}{2}-\delta}$ for $t_i \le T^*$ according to Theorem 5.1.

Consider $T^* < t_i \le T$. Assume we have shown that $|L_{t_j} - \tilde{L}_{t_j}| \le C_j h^{\frac{1}{2}-\delta}$ for $j < i$. We want to derive C_i such that $|L_{t_i} - \tilde{L}_{t_i}| \le C_i h^{\frac{1}{2}-\delta}$ and where all C_i are bounded independent of h. First, consider an intermediate point $s \in (t_{i-1}, t_i)$, then

$$L_s - \tilde{L}_s \le L_{t_{i-1}} + K(s - t_{i-1}) - \tilde{L}_{t_{i-1}} \le L_{t_{i-1}} - \tilde{L}_{t_{i-1}} + Kh,$$

with K the Lipschitz constant of L, and thus

$$\sup_{s\le t_i} |L_s - \tilde{L}_s| \le \max\left(\max_{j<i} |L_{t_j} - \tilde{L}_{t_j}| + Kh, |L_{t_i} - \tilde{L}_{t_i}|\right). \tag{5.33}$$

Now we show that $|L_{t_i} - \tilde{L}_{t_i}| \le C_{t_i} h^{\frac{1}{2}-\delta}$. Consider

$$l_t = \begin{cases} L_t, & t \le T^*, \\ \tilde{L}_t, & t > T^*, \end{cases}$$

and $X^l_t = X_0 + W_t - \alpha l_t$. Then,

$$\begin{aligned}&\left|\mathbb{P}\left(\inf_{s\le t_i} \tilde{X}_s > 0\right) - \mathbb{P}\left(\inf_{s\le t_i} X_s > 0\right)\right| \\ &\le \left|\mathbb{P}\left(\inf_{s\le t_i} \tilde{X}_s > 0\right) - \mathbb{P}\left(\inf_{s\le t_i} X^l_s > 0\right)\right| + \left|\mathbb{P}\left(\inf_{s\le t_i} X^l_s > 0\right) - \mathbb{P}\left(\inf_{s\le t_i} X_s > 0\right)\right|.\end{aligned} \tag{5.34}$$

To estimate the first term, we can write

$$\begin{aligned}\left|\mathbb{P}\left(\inf_{s\le t_i}\tilde{X}_s>0\right)-\mathbb{P}\left(\inf_{s\le t_i}X^l_s>0\right)\right| &= \mathbb{E}\left[\mathbb{1}_{\{\exists t\in[0,t_i]:X^l_t\le 0,\forall s\in[0,t_i]\tilde{X}_s>0\}}\right]\\ &= \mathbb{E}\left[\mathbb{1}_{\{\exists t\in[0,T^*]:X^l_t\le 0,\forall s\in[0,t_i]\tilde{X}_s>0\}}\right]\\ &\le \mathbb{E}\left[\mathbb{1}_{\{\exists t\in[0,T^*]:Y^l_t\le 0,\forall s\in[0,T^*]\tilde{Y}_s>0\}}\right]\\ &= \left|\mathbb{P}\left(\inf_{s\le T^*}\tilde{X}_s>0\right)-\mathbb{P}\left(\inf_{s\le T^*}X^l_s>0\right)\right|,\end{aligned}$$

where we have used in the second line that since $l_t=\tilde{L}_t$ for $t>T^*$, hitting after T^* does not affect the difference. For $t\le T^*$, $l_t=L_t$. Then, using Theorem 5.1,

$$\left|\mathbb{P}\left(\inf_{s\le t_i}\tilde{X}_s>0\right)-\mathbb{P}\left(\inf_{s\le t_i}X^l_s>0\right)\right|\le\left|\mathbb{P}\left(\inf_{s\le T^*}\tilde{X}_s>0\right)-\mathbb{P}\left(\inf_{s\le T^*}X_s>0\right)\right|\le\bar{C}^{T^*}h^{\frac{1}{2}-\delta}.$$

For the second term in (5.34), L_t is Lipschitz on $[T^*,T]$ with $K=\hat{B}(T^*)^{-\frac{1-\beta}{2}}$ and we can apply (5.30). Thus,

$$\left|\mathbb{P}\left(\inf_{s\le t_i}\tilde{X}_s>0\right)-\mathbb{P}\left(\inf_{s\le t_i}X_s>0\right)\right|\le 2\alpha\left(K+\frac{1}{\sqrt{t_i}}\right)\sup_{s\le t_i}|L_s-\tilde{L}_s|+\bar{C}^{T^*}h^{\frac{1}{2}-\delta} \quad (5.35)$$

$$\le 2\alpha\left(K+\frac{1}{\sqrt{t_i}}\right)\max\left(\max_{j<i}|L_{t_j}-\tilde{L}_{t_j}|+Kh,|L_{t_i}-\tilde{L}_{t_i}|\right)+\bar{C}^{T^*}h^{\frac{1}{2}-\delta}, \quad (5.36)$$

using (5.33). Moreover, by (5.24) and (5.36), (5.32) can be written as

$$|L_{t_i}-\tilde{L}_{t_i}|\le 2\alpha\left(K+\frac{1}{\sqrt{t_i}}\right)\max\left(\max_{j<i}|L_{t_j}-\tilde{L}_{t_j}|+Kh,|L_{t_i}-\tilde{L}_{t_i}|\right)+\bar{\gamma}h^{\frac{1}{2}-\delta},$$

where $\bar{\gamma}=\tilde{\gamma}+\bar{C}^{T^*}$.

Taking into account (5.31), we have

$$|L_{t_i}-\tilde{L}_{t_i}|\le\max\left(\left(2\alpha\left(K+\frac{1}{\sqrt{t_i}}\right)\max_{j<i}C_j+\bar{\gamma}+\tilde{\varepsilon}\right)h^{\frac{1}{2}-\delta},\frac{\bar{\gamma}}{1-2\alpha\left(K+\frac{1}{\sqrt{t_i}}\right)}h^{\frac{1}{2}-\delta}\right),$$

where $\tilde{\varepsilon}=2\alpha\left(K+\frac{1}{\sqrt{T^*}}\right)Kh^{\frac{1}{2}+\delta}$.

As a result, we have the following inequality for C_i,

$$C_i\le\max\left(\left(2\alpha\left(K+\frac{1}{\sqrt{t_i}}\right)\max_{j<i}C_j+\bar{\gamma}+\tilde{\varepsilon}\right),\frac{\bar{\gamma}}{1-2\alpha\left(K+\frac{1}{\sqrt{t_i}}\right)}\right).$$

Since the Lipschitz constant $K\le\hat{B}(T^*)^{-\frac{1-\beta}{2}}$, using (5.31), $2\alpha\left(K+\frac{1}{\sqrt{t_i}}\right)<1$. Hence, C_i is bounded independent of h, and by induction we get the result.

□

5.3.3 Monte Carlo simulation of discretized process

In this section, we prove the convergence in probability of

$$\hat{L}_{t_i} = \frac{1}{N}\sum_{k=1}^{N} \mathbb{1}_{\{\min_{j<i} \hat{X}_{t_j}^{(k)} \leq 0\}}$$

in Algorithm 4 to $\tilde{L}_{t_i}$ as $N \to \infty$. We note that we cannot directly apply the law of large numbers, as the summands are dependent through $\hat{L}_{t_j}^N, j < i$. However, we see below that the dependence diminishes (i.e., the covariance goes to zero) as $N \to \infty$, which easily gives convergence, albeit without a Central Limit Theorem-type error estimate or a rate for the variance.

First, we formulate an auxiliary lemma.

Lemma 5.5. *Consider $i \leq n$. Assume for all $j < i$*

$$\hat{L}_{t_j}^N \xrightarrow{\mathbb{P}} \tilde{L}_{t_j}. \tag{5.37}$$

Then,

$$\mathbb{E}[\hat{L}_{t_i}^N] \xrightarrow[N\to\infty]{} \tilde{L}_{t_i} \tag{5.38}$$

$$\mathbb{V}[\hat{L}_{t_i}^N] \xrightarrow[N\to\infty]{} 0. \tag{5.39}$$

Proof. We can write

$$\mathbb{E}[\mathbb{1}_{\{\min_{j<i} \hat{X}_{t_j}^{(k)} > 0\}}] = \mathbb{E}[\mathbb{1}_{\{\min_{j<i}(X_0^{(k)} + W_{t_j}^{(k)} - \alpha \tilde{L}_{t_j} + \alpha(\tilde{L}_{t_j} - \hat{L}_{t_j}^N)) > 0\}}],$$

and estimate

$$\mathbb{E}[\mathbb{1}_{\{\min_{j<i} \tilde{X}_{t_j}^{(k)} > -\alpha \min_{j<i}(\tilde{L}_{t_j} - \hat{L}_{t_j}^N)\}}] \leq \mathbb{E}[\mathbb{1}_{\{\min_{j<i} \hat{X}_{t_j}^{(k)} > 0\}}] \leq \mathbb{E}[\mathbb{1}_{\{\min_{j<i} \tilde{X}_{t_j}^{(k)} > -\alpha \max_{j<i}(\tilde{L}_{t_j} - \hat{L}_{t_j}^N)\}}]. \tag{5.40}$$

We evaluate left- and right-hand side of the last equation separately. We start with the right-hand side, and define for brevity $\tilde{A}_i^{(k)} = \{\min_{j<i} \tilde{X}_{t_j}^{(k)} > -\alpha \max_{j<i}(\tilde{L}_{t_j} - \hat{L}_{t_j}^N)\}$. Consider some $\varepsilon > 0$. Then,

$$\begin{aligned}
\mathbb{E}[\mathbb{1}_{\tilde{A}_j^{(k)}}] = \mathbb{E}[\mathbb{1}_{\tilde{A}_j^{(k)} \cap (\{\max_{j<i}(\tilde{L}_{t_j} - \hat{L}_{t_j}^N) \leq \varepsilon\} \cup \{\max_{j<i}(\tilde{L}_{t_j} - \hat{L}_{t_j}^N) > \varepsilon\})}] = \\
\mathbb{E}[\mathbb{1}_{\tilde{A}_j^{(k)} \cap \{\max_{j<i}(\tilde{L}_{t_j} - \hat{L}_{t_j}^N) \leq \varepsilon\}}] + \mathbb{E}[\mathbb{1}_{\tilde{A}_j^{(k)} \cap \{\max_{j<i}(\tilde{L}_{t_j} - \hat{L}_{t_j}^N) > \varepsilon\}}] \leq \\
\mathbb{P}\left(\tilde{A}_j^{(k)} \cap (\max_{j<i}(\tilde{L}_{t_j} - \hat{L}_{t_j}^N) \leq \varepsilon)\right) + \mathbb{P}\left(\max_{j<i}(\tilde{L}_{t_j} - \hat{L}_{t_j}^N) > \varepsilon\right) \leq \\
\mathbb{P}\left(\min_{j<i} \tilde{X}_{t_j}^{(k)} > -\alpha\varepsilon\right) + \mathbb{P}\left(\max_{j<i}(\tilde{L}_{t_j} - \hat{L}_{t_j}^N) > \varepsilon\right) = \\
\mathbb{P}\left(\min_{j<i} \tilde{X}_{t_j}^{(k)} > 0\right) + \left(\mathbb{P}\left(\min_{j<i} \tilde{X}_{t_j}^{(k)} > -\alpha\varepsilon\right) - \mathbb{P}\left(\min_{j<i} \tilde{X}_{t_j}^{(k)} > 0\right)\right) + \mathbb{P}\left(\max_{j<i}(\tilde{L}_{t_j} - \hat{L}_{t_j}^N) > \varepsilon\right) \leq \\
1 - \tilde{L}_{t_i} + \alpha\varepsilon \sup_{\theta\in[0,1]} \tilde{\varphi}_i(-\theta\alpha\varepsilon) + \mathbb{P}\left(\max_{j<i}(\tilde{L}_{t_j} - \hat{L}_{t_j}^N) > \varepsilon\right),
\end{aligned}$$

where $\tilde{\varphi}_i(x)$ is the pdf of $\min_{j<i} \tilde{X}^{(k)}_{t_j}$, which is bounded according to Lemma 5.3.

Using (5.37) and properties of convergence in probability, we have

$$\max_{j<i}(\tilde{L}_{t_j} - \hat{L}^N_{t_j}) \xrightarrow{\mathbb{P}} 0.$$

Thus, the last term goes to 0 with $N \to \infty$. Considering $\varepsilon \to 0$, we get

$$\lim_{N\to\infty} \mathbb{E}[\mathbb{1}_{\tilde{A}^{(k)}_j}] \leq 1 - \tilde{L}_{t_i}. \tag{5.41}$$

Similarly, denote by $\bar{A}^{(k)}_i = \{\min_{j<i} \tilde{X}^{(k)}_{t_j} > -\alpha \min_{j<i}(\tilde{L}_{t_j} - \hat{L}^N_{t_j})\}$, then

$$\begin{aligned}
\mathbb{E}[\mathbb{1}_{\bar{A}^{(k)}_j}] = \mathbb{E}[\mathbb{1}_{\tilde{A}^{(k)}_j \cap (\{\min_{j<i}(\tilde{L}_{t_j} - \hat{L}^N_{t_j})\} \geq -\varepsilon\} \cup \{\min_{j<i}(\tilde{L}_{t_j} - \hat{L}^N_{t_j})\} < -\varepsilon\})}] = \\
\mathbb{E}[\mathbb{1}_{\bar{A}^{(k)}_j \cap \{\min_{j<i}(\tilde{L}_{t_j} - \hat{L}^N_{t_j}) \geq -\varepsilon\}}] + \mathbb{E}[\mathbb{1}_{\bar{A}^{(k)}_j \cap \{\min_{j<i}(\tilde{L}_{t_j} - \hat{L}^N_{t_j}) < -\varepsilon\}}] \geq \\
\mathbb{P}\left(\bar{A}^{(k)}_j \cap \{\min_{j<i}(\tilde{L}_{t_j} - \hat{L}^N_{t_j}) \geq -\varepsilon\}\right) \geq \\
\mathbb{P}\left(\min_{j<i} \tilde{X}^{(k)}_{t_j} > \alpha\varepsilon\right) \geq 1 - \tilde{L}_{t_i} - \alpha\varepsilon \inf_{\theta\in[0,1]} \tilde{\varphi}_i\left(\theta\alpha\varepsilon\right),
\end{aligned}$$

where $\tilde{\varphi}_i(x)$ is the pdf of $\min_{j<i} \tilde{X}^{(k)}_{t_j}$, which is bounded according to Lemma 5.3. Thus,

$$\lim_{N\to\infty} \mathbb{E}[\mathbb{1}_{\tilde{A}^{(k)}_j}] \geq 1 - \tilde{L}_{t_i}. \tag{5.42}$$

Combining (5.41), (5.42), and (5.40), we immediately get (5.38).

Consider now the variance

$$\mathbb{V}[\hat{L}^N_{t_i}] = \frac{1}{N^2}\left(\sum_{k=1}^{N} \mathbb{V}[\mathbb{1}_{\{\min_{j<i} \hat{X}^{(k)}_{t_j} > 0\}}] + \sum_{k\neq l} \operatorname{cov}\left(\mathbb{1}_{\{\min_{j<i} \hat{X}^{(k)}_{t_j} > 0\}}, \mathbb{1}_{\{\min_{j<i} \hat{X}^{(l)}_{t_j} > 0\}}\right)\right). \tag{5.43}$$

The covariance can be estimated by

$$\begin{aligned}
\operatorname{cov}\left(\mathbb{1}_{\{\min_{j<i} \hat{X}^{(k)}_{t_j} > 0\}}, \mathbb{1}_{\{\min_{j<i} \hat{X}^{(l)}_{t_j} > 0\}}\right) = \\
\mathbb{E}\left[\mathbb{1}_{\{\min_{j<i} \hat{X}^{(k)}_{t_j} > 0\}} \mathbb{1}_{\{\min_{j<i} \hat{X}^{(l)}_{t_j} > 0\}}\right] - \mathbb{E}\left[\mathbb{1}_{\{\min_{j<i} \hat{X}^{(k)}_{t_j} > 0\}}\right] \mathbb{E}\left[\mathbb{1}_{\{\min_{j<i} \hat{X}^{(l)}_{t_j} > 0\}}\right].
\end{aligned} \tag{5.44}$$

As above and because of exchangeability, for the second term of the last expression,

$$\mathbb{E}\left[\mathbb{1}_{\{\min_{j<i} \hat{X}^{(k)}_{t_j} > 0\}}\right] \mathbb{E}\left[\mathbb{1}_{\{\min_{j<i} \hat{X}^{(l)}_{t_j} > 0\}}\right] \xrightarrow[N\to\infty]{} \left(1 - \tilde{L}_{t_i}\right)^2. \tag{5.45}$$

Now we consider the first term of (5.44). Similar to above, it can be estimated by

$$\begin{aligned}
\mathbb{E}[\mathbb{1}_{\{\min_{j<i} \tilde{X}^{(k)}_{t_j} > -\alpha \min_{j<i}(\tilde{L}_{t_j} - \hat{L}^N_{t_j})\}} \mathbb{1}_{\{\min_{j<i} \tilde{X}^{(l)}_{t_j} > -\alpha \min_{j<i}(\tilde{L}_{t_j} - \hat{L}^N_{t_j})\}}] \leq \\
\mathbb{E}\left[\mathbb{1}_{\{\min_{j<i} \hat{X}^{(k)}_{t_j} > 0\}} \mathbb{1}_{\{\min_{j<i} \hat{X}^{(l)}_{t_j} > 0\}}\right] \\
\leq \mathbb{E}[\mathbb{1}_{\{\min_{j<i} \tilde{X}^{(k)}_{t_j} > -\alpha \max_{j<i}(\tilde{L}_{t_j} - \hat{L}^N_{t_j})\}} \mathbb{1}_{\{\min_{j<i} \tilde{X}^{(l)}_{t_j} > -\alpha \max_{j<i}(\tilde{L}_{t_j} - \hat{L}^N_{t_j})\}}].
\end{aligned} \tag{5.46}$$

Similar to (5.41) and (5.42), one can show that

$$\lim_{N\to\infty} \mathbb{E}[\mathbb{1}_{\{\min_{j<i} \tilde{X}^{(k)}_{t_j} \ > -\alpha \min_{j<i}(\tilde{L}_{t_j}-\hat{L}^N_{t_j})\}} \mathbb{1}_{\{\min_{j<i} \tilde{X}^{(l)}_{t_j} \ > -\alpha \min_{j<i}(\tilde{L}_{t_j}-\hat{L}^N_{t_j})\}}] \geq (1-\tilde{L}_{t_i})^2,$$
$$\lim_{N\to\infty} \mathbb{E}[\mathbb{1}_{\{\min_{j<i} \tilde{X}^{(k)}_{t_j} \ > -\alpha \max_{j<i}(\tilde{L}_{t_j}-\hat{L}^N_{t_j})\}} \mathbb{1}_{\{\min_{j<i} \tilde{X}^{(l)}_{t_j} \ > -\alpha \max_{j<i}(\tilde{L}_{t_j}-\hat{L}^N_{t_j})\}}] \leq (1-\tilde{L}_{t_i})^2.$$

Hence,

$$\mathbb{E}\left[\mathbb{1}_{\{\min_{j<i} \hat{X}^{(k)}_{t_j}>0\}} \mathbb{1}_{\{\min_{j<i} \hat{X}^{(l)}_{t_j}>0\}}\right] \xrightarrow[N\to\infty]{} (1-\tilde{L}_{t_i})^2. \tag{5.47}$$

Thus, combining (5.45) and (5.47), we get

$$\operatorname{cov}\left(\mathbb{1}_{\{\min_{j<i} \hat{X}^{(k)}_{t_j}>0\}}, \mathbb{1}_{\{\min_{j<i} \hat{X}^{(l)}_{t_j}>0\}}\right) \xrightarrow[N\to\infty]{} 0.$$

Considering the first term of (5.43), it is easy to see that each variance is bounded by 1. As a result, we have

$$\mathbb{V}[\hat{L}^N_{t_i}] \leq \frac{1}{N} + \max_{k\neq l} \left|\operatorname{cov}\left(\mathbb{1}_{\{\min_{j<i} \hat{X}^{(k)}_{t_j}>0\}}, \mathbb{1}_{\{\min_{j<i} \hat{X}^{(l)}_{t_j}>0\}}\right)\right| \xrightarrow[N\to\infty]{} 0.$$

□

Now we can deduce the convergence instantly.

Proof of Theorem 5.2. The proof is immediate by induction. The statement is true for $i = 0$. Now take $i \geq 1$. By Lemma 5.5, there exists N^* such that for all $N > N^*$,

$$|\mathbb{E}[\hat{L}^N_{t_i}] - \tilde{L}_{t_i}| \leq \frac{\varepsilon}{2}.$$

Thus, by Chebyshev's inequality, we have

$$\mathbb{P}(|\hat{L}^N_{t_i} - \tilde{L}_{t_i}| > \varepsilon) \leq \mathbb{P}\left(|\hat{L}^N_{t_i} - \mathbb{E}[\hat{L}^N_{t_i}]| > \frac{\varepsilon}{2}\right) \leq \frac{4\mathbb{V}[\hat{L}^N_{t_i}]}{\varepsilon^2}.$$

Using again Lemma 5.5, we have that $\mathbb{V}[\hat{L}^N_{t_i}] \xrightarrow[N\to\infty]{} 0$. Hence,

$$\hat{L}^N_{t_i} \xrightarrow[N\to\infty]{\mathbb{P}} \tilde{L}_{t_i},$$

for i and by induction we have proved the theorem. □

5.3.4 Brownian bridge convergence improvement

In this section, we prove Theorem 5.3, which ascertains the uniform convergence (in t) of $\breve{L}$ to L at the improved rate.

Proof of Theorem 5.3. We shall proceed by induction. For $i = 0$, we have that $\breve{L}_0 = L_0 = 0$. Assume we have shown that $|\breve{L}_{t_j} - L_{t_j}| \leq C_j h^{\frac{1+\beta}{2}}$ for all $j < i$ with some $C_j > 0$, and we want to estimate $|\breve{L}_{t_i} - L_{t_i}|$. First, for $j > 0$ we have

$$\sup_{t_j \leq s < t_{j+1}} |\breve{L}_s - L_s| \leq |\breve{L}_{t_j} - L_{t_j}| + \hat{B} h^{\frac{1+\beta}{2}} \leq (C_j + \hat{B}) h^{\frac{1+\beta}{2}}, \tag{5.48}$$

since $L'_\zeta \le \hat{B}\zeta^{-\frac{1-\beta}{2}}$, while for $j=0$, $\sup_{0\le s<t_1}|\breve{L}_s - L_s| \le L_{t_1} \le \tilde{B}h^{\frac{1+\beta}{2}}$. Now consider

$$\begin{aligned}
|\breve{L}_{t_i} - L_{t_i}| &= \left|\mathbb{P}\left(\inf_{s\le t_i}\breve{X}_s > 0\right) - \mathbb{P}\left(\inf_{s\le t_i} X_s > 0\right)\right| \\
&= \left|\mathbb{P}\left(\inf_{s\le t_i}(X_s + \alpha(L_s - \breve{L}_s)) > 0\right) - \mathbb{P}\left(\inf_{s\le t_i} X_s > 0\right)\right| \\
&\le \left|\mathbb{P}\left(\inf_{s\le t_i} X_s > -\alpha\sup_{s<t_i}|\breve{L}_s - L_s|\right) - \mathbb{P}\left(\inf_{s\le t_i} X_s > 0\right)\right| \\
&\le \mathbb{P}\left(\inf_{s\le t_i} X_s > -\alpha\max_{j<i}(C_j + \hat{B})h^{\frac{1+\beta}{2}}\right) - \mathbb{P}\left(\inf_{s\le t_i} X_s > 0\right) \\
&\le \alpha\max_{j<i}(C_j + \hat{B})\sup_{\theta\in[0,1]}\varphi\left(-\theta\alpha\max_{j<i}(C_j+\hat{B})h^{\frac{1+\beta}{2}}\right)h^{\frac{1+\beta}{2}},
\end{aligned}$$

where $\varphi_i(x)$ is the density of $\inf_{s<t_i} X_s$.

Then, using Lemma 5.3, we have

$$C_i \le \alpha B\max_{j<i}(C_j+\hat{B})\left[\sqrt{\frac{2t_i}{\pi}} + \alpha\tilde{B}t_i^{\frac{1+\beta}{2}}\right]^\beta \le \gamma\sum_{k=0}^{i}(\alpha B)^k\prod_{j=1}^{k}\left[\sqrt{\frac{2t_j}{\pi}} + \alpha\tilde{B}t_j^{\frac{1+\beta}{2}}\right]^\beta,$$

where $\gamma = \alpha B\hat{B}\left[\sqrt{\frac{2T}{\pi}} + \alpha\tilde{B}T^{\frac{1+\beta}{2}}\right]^\beta$.

Thus, C_i is bounded independent of h and i by (5.10). By induction we get (5.20).

□

The proof of Theorem 5.3 indicates that the order is limited by the behaviour of L for small t. The next result shows that a locally refined time mesh achieves convergence order 1 for all β.

Corollary 5.1. *Consider a non-uniform time mesh $t_i = (ih)^{\frac{2}{1+\beta}}$ for $0 \le i \le n$ with $h = T^{\frac{1+\beta}{2}}/n$. Then, there exists $C_1 > 0$, independent of h, such that*

$$\max_{i\le n}|\breve{L}_{t_i} - L_{t_i}| \le C_1 h.$$

Proof. We shall proceed by induction as in the proof of Theorem 5.3 above. We now have the time step

$$t_{j+1} - t_j = ((j+1)h)^{\frac{2}{1+\beta}} - (jh)^{\frac{2}{1+\beta}} = \frac{2}{1+\beta}h^{\frac{2}{1+\beta}}(j+\xi)^{\frac{1-\beta}{1+\beta}},$$

for some $\xi \in (0,1)$, and, for $j > 0$,

$$L'_\zeta \le \hat{B}t_j^{-\frac{1-\beta}{2}} = \hat{B}(jh)^{-\frac{1-\beta}{1+\beta}}.$$

Hence, for $j > 0$,

$$\sup_{t_j\le s<t_{j+1}}|\breve{L}_s - L_s| \le |\breve{L}_{t_j} - L_{t_j}| + L'_\zeta(t_{j+1} - t_j) \le |\breve{L}_{t_j} - L_{t_j}| + \bar{B}h, \tag{5.49}$$

for some $\bar{B} > 0$ independent of h and j. We treat $j = 0$ separately and obtain directly

$$\sup_{0\le s<t_1}|\breve{L}_s - L_s| \le L_{t_1} \le \tilde{B}h.$$

Repeating the remaining steps of the proof of Theorem 5.3, we get the result.

□

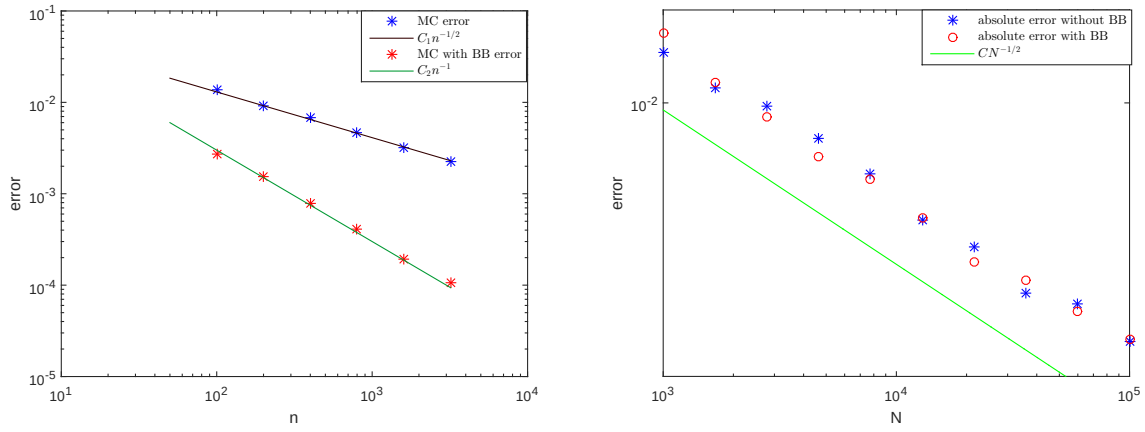
5.4 Numerical experiments

In this section, we demonstrate that the proven convergence orders in the timestep of $1/2$, $(1+\beta)/2$, and 1 for the different methods are indeed sharp in the case of regular solutions; that the empirical variance of N-sample estimators is $1/N$; and that the method also converges experimentally in the presence of blow-up. To show this, we study three test cases with varying regularity of the initial data and of the loss function.

5.4.1 Lipschitz initial data and no blow-up

In our first experiment, we choose X_0 such that $\frac{1}{X_0} \sim \exp(\lambda)$, which guarantees that the density decays exponentially near zero. We take $\lambda = 1$ in our experiments, and pick the parameters in (5.1) to be $\alpha = 0.8$, $T = 2$. The solution is found to be continuous.

We perform numerical simulations using Algorithms 4 and 5 with $N = 2 \times 10^5$ particles and different time meshes varying from 50 to 3200 points. To estimate the error, we consider the difference $|\tilde{L}_T^{2n} - \tilde{L}_T^n|$ between the solutions with n and $2n$ timesteps, respectively, computed with the same paths. The results, presented in Figure 5.2, agree with the theory: for Algorithm 4, we get the convergence rate $\frac{1}{2}$, and for Algorithm 5, the rate is 1, because the initial distribution is regular enough around 0, i.e. $\beta = 1$ in (5.1). We also investigate the convergence rate of $\hat{L}_{t_i}$ to $\tilde{L}_{t_i}$ and $\bar{L}_{t_i}$ to $\breve{L}_{t_i}$ empirically.

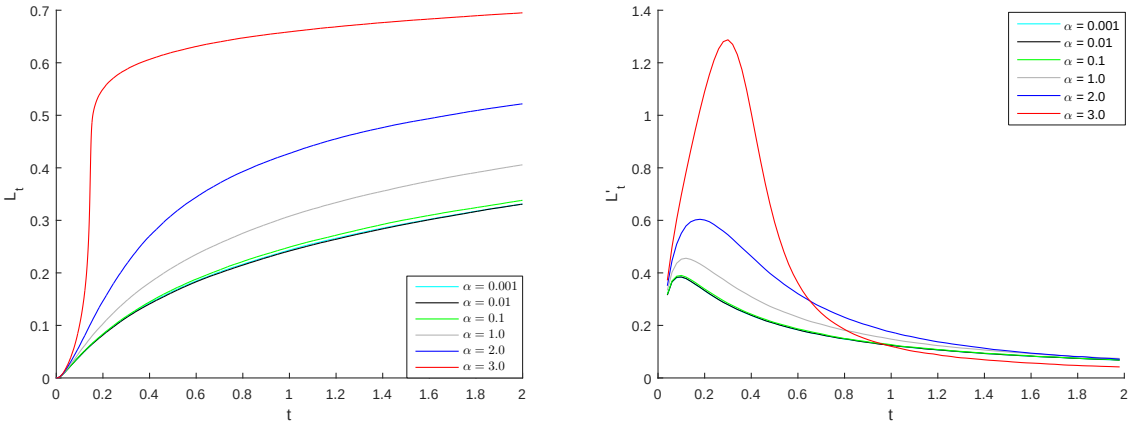


(a) (b)

Figure 5.2: Error of the loss process at $t = T$ for $\frac{1}{X_0} \sim \exp(1)$: (a) for increasing number n of timesteps; (b) for increasing number N of samples, both for Algorithms 4 and 5.

In order to compute the benchmark solution, we used $N = 5 \times 10^7$ particles. From the results we conclude that both Algorithm 4 and 5 have the convergence rate $\frac{1}{2}$ in N.

To illustrate the dependence on the parameter α, we include in Figure 5.3 plots for L_t and L_t' for different values of α. We evaluate L_t' numerically using a central finite difference approximation. In order to reduce the Monte Carlo noise, we increase N to 5×10^7 and reduce n to 200.



(a) (b)

Figure 5.3: (a) L_t and (b) L_t' for different values of α.

5.4.2 Hölder $1/2$ initial data and no blow-up

In another example, we again consider $\alpha = 0.8$ and $T = 2$, but choose $X_0 \sim \text{Gamma}(1+\beta, 1/2)$, i.e. with density

$$f_{X_0}(x) = \frac{x^\beta \mathrm{e}^{-x/2}}{\Gamma(1+\beta)2^{1+\beta}},$$

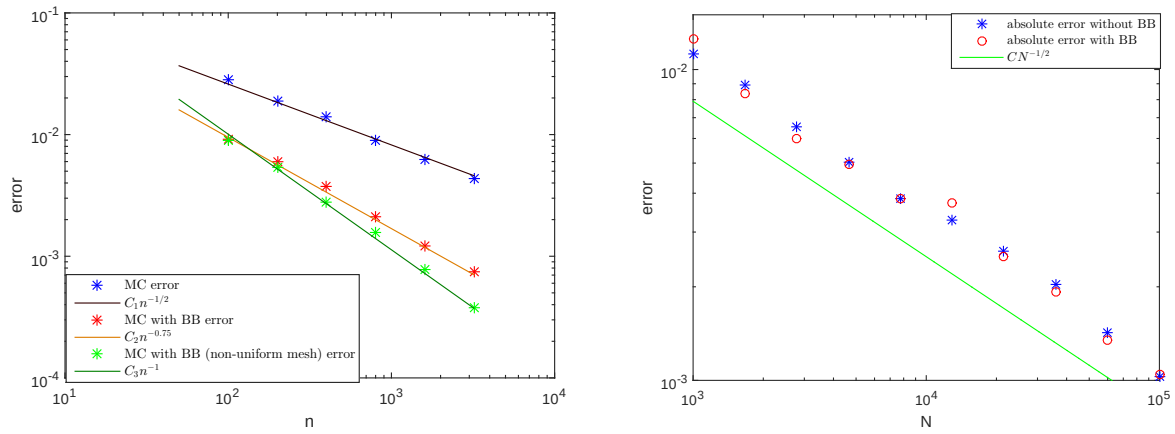
such that we have that $f_{Y_0}(x) \leq Cx^\beta$ for $x > 0$ and some $C > 0$. We choose $\beta = \frac{1}{2}$. The solution is again found to be continuous.

We perform numerical simulations using Algorithm 4, and Algorithm 5 on uniform and non-uniform meshes varying from 50 to 3200 points and with $N = 2 \times 10^5$ particles. The results are presented in Figure 5.4. As predicted by the theory, for Algorithm 4 we get the convergence rate $\frac{1}{2}$; for Algorithm 5 on uniform meshes, we get the rate $\frac{1+\beta}{2} = \frac{3}{4}$; and for Algorithm 5 on non-uniform meshes, we get the rate 1. As in the previous example, we also investigate the convergence rate in N empirically. These results also confirm $\frac{1}{2}$ convergence rate in N for both Algorithms 4 and 5. In Figure 5.5 we present the dependence of L_t and L_t' on the parameter α. As in the previous example, we use $N = 5 \times 10^7$ and $n = 200$.

5.4.3 Hölder $1/2$ initial data and blow-up

In our third example, we illustrate possible jumps arising in the solution for sufficiently large values of α. We consider $\alpha = 1.5$, $T = 0.008$ and choose Y_0 as in the previous example, $X_0 \sim \text{Gamma}(1+\beta, 1/2)$. Note that the blow-up happens already for very small t due to the interplay of the mass close to 0 for X_0 and the relatively large α.

With the lack of convergence theory for discontinuous $(L_t)_{t\geq 0}$, we apply Algorithms 4 and 5, and empirically estimate the error. In Figure 5.6 (a) we show $\tilde{L}_t$ computed using Algorithm 4 for different n; in Figure 5.6 (b) the numerical error as a function of t for specific n; and in Figure 5.7 (a) and (b) we estimate the convergence rate for different t.



(a) (b)

Figure 5.4: Error of the loss process at $t = T$ for $X_0 \sim \text{Gamma}(3/2, 1/2)$: (a) for increasing number n of timesteps; (b) for increasing number N of samples, both for Algorithms 4 and 5.

Figure 5.6 (a) shows that a fairly fine resolution is needed to capture the discontinuity and its timing, but that all meshes predict the size of the jump well. This is further illustrated in Figure 5.6 (b), which shows the build-up of the error before the jump, the lack of uniform convergence due to the displacement of the jump on different meshes, and the relatively constant error after the jump.

In Figure 5.7 (a) we estimate the convergence order at $T = 0.002$, i.e. before the jump. By regression, we get 0.53 $(0.47, 0.59)$ for Algorithm 4 and 0.80 $(0.72, 0.88)$ for Algorithm 5, where the 95% confidence interval is in brackets, which agrees with the theory for continuous L_t. In Figure 5.7 (b), for the error at $T = 0.008$, i.e. after the jump, we get 0.93 $(0.81, 1.05)$ for Algorithm 4 and 1.03 $(0.92, 1.14)$ for Algorithm 5. The faster convergence may result from the almost constancy of the losses after the jump.

To get an overall picture of the accuracy, we suggest the following metrics to measure the "closeness" of the computed solutions to L_t:

1. $d_1(L_t, \tilde{L}_t) = \int_0^T |L_t - \tilde{L}_t| \, \mathrm{d}t$. This is practically computed by numerical integration.

2. $d_2(L_t, \tilde{L}_t) = |t_* - \tilde{t}_*|$, where t_* and $\tilde{t}_*$ are the jump times for L_t and $\tilde{L}_t$, respectively. They are approximated by the points with the steepest gradient $j_* = \text{argmax}_j(\tilde{L}_{t_j} - \tilde{L}_{t_{j-1}}), t_* = j_* h$.

3. $d_3(L_t, \tilde{L}_t) = \sup_{t \in [0,T]} |L_t^{-1} - \tilde{L}_t^{-1}|$, where L_t^{-1} and $\tilde{L}_t^{-1}$ are the corresponding inverse functions. The inverse functions are found by the chebfun toolbox (Driscoll et al. (2014)), which automatically splits functions into intervals of continuity.

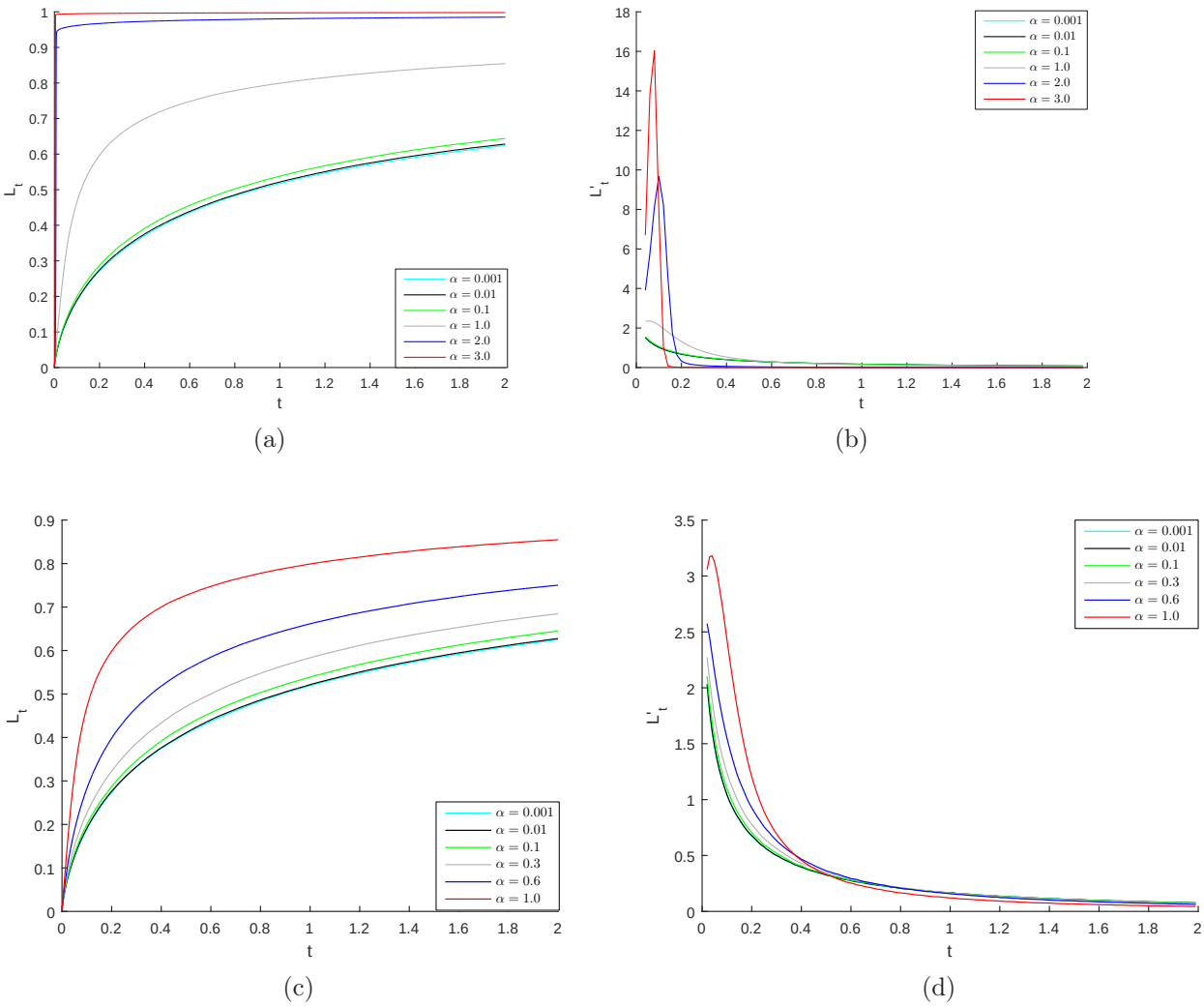
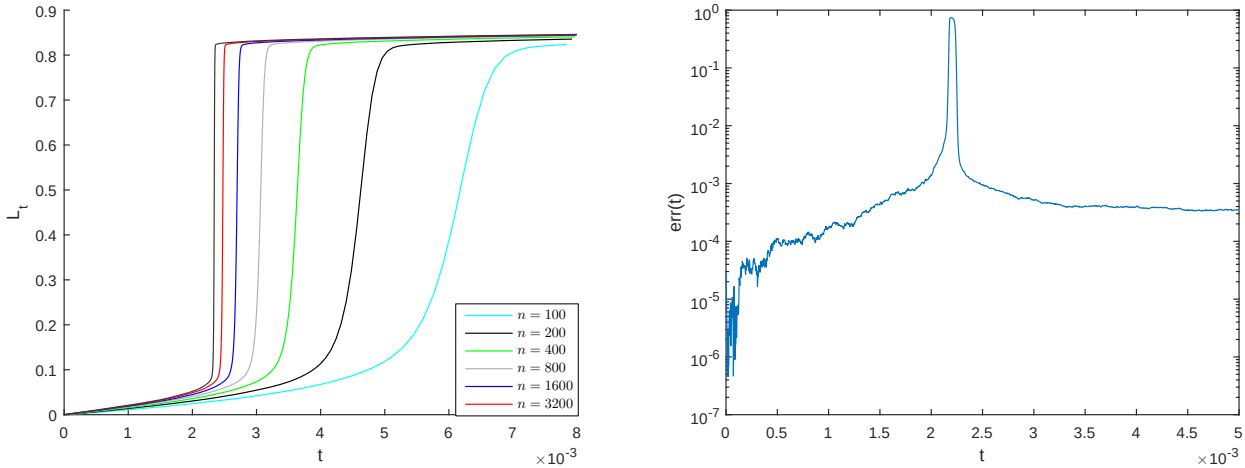


Figure 5.5: (a) L_t and (b) L'_t for different values of α.

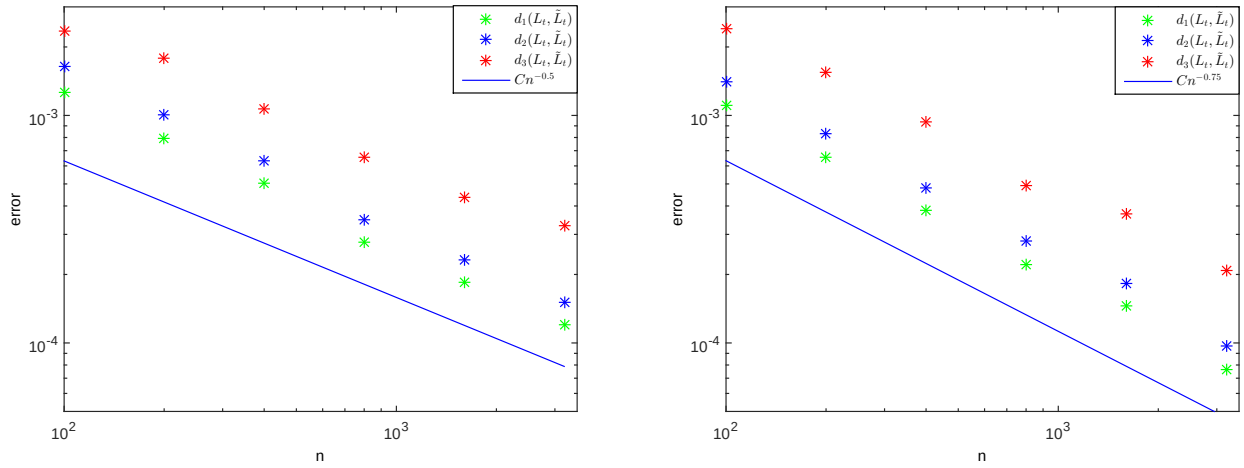
In the absence of the exact L_t, we use again the difference between quantities computed using $2n$ and n points, for the same Monte Carlo paths, as a proxy.

We present the results in Figure 5.8 (a) for Algorithm 4 and in Figure 5.8 (b) for Algorithm 5. For metric 1 we have convergence rate 0.68 $(0.63, 0.73)$ and 0.76 $(0.71, 0.81)$, for metric 2 we have 0.70 $(0.65, 0.80)$ and 0.76 $(0.72, 0.80)$, and for metric 3 we have 0.60 $(0.52, 0.68)$ and 0.71 $(0.63, 0.79)$ for Algorithms 4 and 5, respectively, where the 95% confidence interval are in brackets. We observe that the convergence rate for Algorithm 4 is somewhat better than 0.5, while for Algorithm 5 it is around 0.75 for all three metrics.



(a) (b)

Figure 5.6: (a) Loss process computed using Algorithm 4 for different n; (b) error as a function of t.



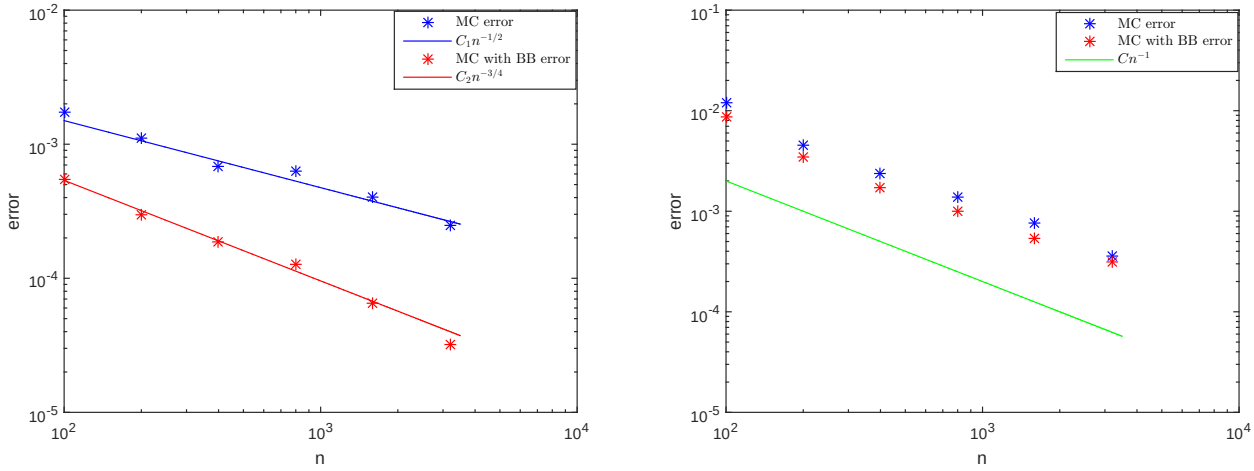
(a) (b)

Figure 5.8: Convergence rate for $d_1(L_t, \tilde{L}_t)$, $d_2(L_t, \tilde{L}_t)$, $d_3(L_t, \tilde{L}_t)$ for (a) Algorithm 4, (b) Algorithm 5.

5.5 Conclusion

We have developed particle methods with explicit timestepping for the simulation of (5.1)-(5.3). Convergence with a rate up to 1 in the timestep is shown under a condition on the model parameters and time horizon, when the loss function is differentiable. Experimentally, the method also converges in the blow-up regime, and the variance of estimators is inversely proportional to the number of samples.

This opens up several theoretical and practical questions. The efficiency of the method could be significantly improved by a simple application of multilevel simulation (Szpruch



(a) (b)

Figure 5.7: Convergence rate: at (a) $T = 0.001$, (b) $T = 0.008$.

et al. (2017); Ricketson (2015)). If the variance can be shown to behave on the finer levels as suggested by the numerical tests, combined with the proven result on the time stepping bias, the computational complexity for root-mean-square error ϵ would be brought down to ϵ^{-2}, from ϵ^{-3} as observed presently (ϵ^{-2} from the number of samples and ϵ^{-1} from the number of timesteps).

For the particular system (5.1)–(5.3), it is conceivable to apply a particle method without time stepping of the form

$$X_t^{(i)} \quad = \quad X_0^{(i)} + W_t^{(i)} - \alpha \frac{1}{N} \sum_{j=1}^{N} \mathbb{1}_{\{\tau_{(j)} \leq t\}},$$

where $X_0^{(i)}$ and $W^{(i)}$ are N i.i.d. copies of Y_0 and W, and $\tau_{(j)}$, $1 \leq j \leq N$, is the order statistic of the hitting times of zero. Then, for $\tau_{(i)} < t < \tau_{(i+1)}$, conditional on the information up to $\tau_{(i)}$, the law of $X_t^{(j)} - X_{\tau_{(i)}}^{(j)}$ for those $N - i$ particles j which have not yet hit, is identical to that of independent standard Brownian motions with constant drift $-\alpha i/N$. So if we simulate those independent hitting times of $-X_{\tau_{(i)}}^{(j)}$ for each and then take the smallest one to be $\tau_{(i+1)}$, this inductive construction satisfies the correct law. The complexity is then $O(N^2)$, and combined with the observed variance of $O(1/N)$, this gives a complexity of ϵ^{-4} for error ϵ. Another disadvantage is that this method with exact sampling is restricted to the constant parameter case, while the time stepping scheme can more easily be extended to variable coefficients.

Theoretically, one would like guaranteed convergence also in the blow-up regime. This requires the choice of an appropriate metric – the Skorokhod distance may be suitable, considering the analysis in Delarue et al. (2015b) and the tests in Section 5.4.3. It also requires verification that the limit in the case of blow-up is a so-called "physical solution" as defined in Hambly et al. (2018); Nadtochiy and Shkolnikov (2017); Delarue et al. (2015b), which on the one hand results from a specific sequential realisation of the losses in the case of blow-up (Nadtochiy and Shkolnikov (2017); Delarue et al. (2015b)), and

on the other hand is the right-continuous solution with the smallest jump size (Hambly et al. (2018)).

Chapter 6

Integral equation methods for the infinite-dimensional model

6.1 Introduction

In this chapter we solve (2.80), or more precisely (2.83) using integral equation methods. For simplicity, we assume that the initial state of all banks is the same, i.e. $X_0 = z$ for some $z > 0$. To emphasize the dependence on z, in the following we denote the density as $p(t, x; z)$. Then, (2.83) transforms to

$$\begin{aligned} p_t(t, x; z) &= \frac{\alpha}{2} p_x(t, 0; z)\, p_x(t, x; z) + \frac{1}{2} p_{xx}(t, x; z), \\ p(0, x; z) &= \delta(x - z), \\ p(t, 0; z) &= 0. \end{aligned} \tag{6.1}$$

First, for *given* drift term from the mean-field interaction, the problem is formulated as diffusion problem on a semi-infinite domain with curvilinear boundary, and its solution represented semi-analytically by the method of heat potentials. A detailed introduction to heat potentials can be found in Appendix B.

The first use of the method of heat potentials in mathematical finance is found in Lipton (2001) for pricing path-dependent options with curvilinear barrier (Section 12.2.3, pp. 462–467).

Second, expressing the interaction term by the solution from the first step results in two coupled Volterra equations. For early applications of heat potentials to versions of the Stefan problem see already, e.g., (Rubinstein, 1971, Part II, Chapter 1). These singular Volterra equations are then solved by either an expansion method for small interaction parameter, or numerically by discretisation and Newton-Raphson iteration.

An expansion approach for a certain McKean–Vlasov equation with mean-field interaction through the drift was recently studied in Gobet and Pagliarani (2018), who perform an iterative two-step procedure which decouples the nonlinearity arising from the dependence of the drift on the law of the process from the standard dependence on the state-variables. The present paper differs not only in the solution approach taken, but fundamentally in the considered mean-field interaction (through hitting times rather

than the expectation of the drift) and the parameter of the expansion (for small drift interaction rather than small volatility).

To assess the accuracy of the (first order) perturbation solution and to illustrate the behaviour for strong interactions where the expansion breaks down, we describe a simple numerical algorithm, but refer the reader to the large and well-established body of literature on more advanced numerical methods for Volterra equations (see Section 6.3.3).

The rest of the chapter is organized as follows: in Section 6.2, we derive a solution for the first passage density of Brownian motion over a curve in terms of a Volterra equation, using the method of heat potentials, and thus obtain the interaction term in the original McKean–Vlasov equation; in Section 6.3, we we consider a perturbation method and a numerical method for the corresponding system of Volterra equations; in Section 6.4, we show numerical illustrations and compare the methods; in Section 6.5 we conclude.

6.2 The method of heat potentials

In this section we compute the transition density and the first passage density of Brownian motion with a known time-dependent drift $\mu(t)$ on the positive semi-axis. The transition probability density $p(t,x,z)$ satisfies the Kolmogorov forward equation (2.81).

We first derive an expression for $p(t,x;z)$ using the method of heat potentials with an unknown weight function which can be found as a solution of a Volterra equation of the second kind. Next, we differentiate the expression for $p(t,x;z)$ with respect to x and take its limit to 0 in order to find the first passage density of X in (2.80), or, equivalently, g in (2.82). This limit is less well-known, so we calculate it for completeness.

Below we omit z when possible.

6.2.1 Semi-closed formula for the transition density

Let $M(t) = \int_0^t \mu(t')\,\mathrm{d}t'$. The change of variables $(t,x) \to (t,y) = (t, x - M(t))$ yields the following Cauchy-Dirichlet problem:

$$\begin{aligned} &p_t(t,y) = \frac{1}{2} p_{yy}(t,y), \\ &p(0,y) = \delta(y-z), \\ &p(t,-M(t)) = 0. \end{aligned} \tag{6.2}$$

We split p in two parts:

$$p(t,y) = q(t,y) + H(t,y),$$

where $H(t,y)$ is the standard heat kernel,

$$H(t,y) = \frac{\exp\left(-\frac{(y-z)^2}{2t}\right)}{\sqrt{2\pi t}}.$$

The corresponding problem for q has the form

$$\begin{aligned} q_t(t,y) &= \frac{1}{2} q_{yy}(t,y), \\ q(0,y) &= 0, \\ q(t,-M(t)) &= -\frac{\exp\left(-\frac{(M(t)+z)^2}{2t}\right)}{\sqrt{2\pi t}}. \end{aligned}$$

We use the method of heat potentials (see Lipton (2001), pp. 462–468). Thereby, we represent q in the form

$$q(t,y) = \int_0^t \frac{(y + M(t')) \exp\left(-\frac{(y+M(t'))^2}{2(t-t')}\right)}{\sqrt{2\pi (t-t')^3}} \nu(t')\, \mathrm{d}t',$$

where ν is a suitable weight which will be determined to match the boundary condition. Assuming that $\nu(t)$ is known, we can revert to the original variables and get

$$p(t,x) = \int_0^t \frac{(x - \Psi(t,t')) \exp\left(-\frac{(x-\Psi(t,t'))^2}{2(t-t')}\right)}{\sqrt{2\pi (t-t')^3}} \nu(t')\, \mathrm{d}t' + \frac{\exp\left(-\frac{(x-M(t)-z)^2}{2t}, \right)}{\sqrt{2\pi t}}, \qquad (6.3)$$

where $\Psi(t,t') = M(t) - M(t')$.

By construction, p in (6.3) satisfies the first two equations in (6.2). We also need to satisfy the boundary condition in (6.2). It is easy to show that

$$\begin{aligned} \mathbb{L}_1 &:= \lim_{x \to 0} \int_0^t \frac{(x - \Psi(t,t')) \exp\left(-\frac{(x-\Psi(t,t'))^2}{2(t-t')}\right)}{\sqrt{2\pi (t-t')^3}} \nu(t')\, \mathrm{d}t' \\ &= \nu(t) - \int_0^t \frac{\Psi(t,t')\, \Xi(t,t')}{\sqrt{2\pi (t-t')^3}} \nu(t')\, \mathrm{d}t'. \end{aligned}$$

We split $\mathbb{L}_1$ into two parts,

$$\mathbb{L}_1 = \mathbb{L}_1^{(1)} - \mathbb{L}_1^{(2)},$$

where

$$\begin{aligned} \mathbb{L}_1^{(1)} &= \lim_{x \to 0} \left(x \int_0^t \frac{\exp\left(-\frac{(x-\Psi(t,t'))^2}{2(t-t')}\right)}{\sqrt{2\pi (t-t')^3}} \nu(t')\, \mathrm{d}t' \right), \\ \mathbb{L}_1^{(2)} &= \lim_{x \to 0} \int_0^t \frac{\Psi(t,t') \exp\left(-\frac{(x-\Psi(t,t'))^2}{2(t-t')}\right)}{\sqrt{2\pi (t-t')^3}} \nu(t')\, \mathrm{d}t'. \end{aligned}$$

We represent $\nu(t') = \nu(t) + (\nu(t') - \nu(t))$, and write

$$\begin{aligned}\mathbb{L}_1^{(1)} &= \lim_{x\to 0}\left(x\nu(t)\int_0^t \frac{\exp\left(-\frac{(x-\Psi(t,t'))^2}{2(t-t')}\right)}{\sqrt{2\pi(t-t')^3}}\,\mathrm{d}t'\right)\\ &\quad+\lim_{x\to 0}\left(x\int_0^t \frac{\exp\left(-\frac{(x-\Psi(t,t'))^2}{2(t-t')}\right)}{\sqrt{2\pi(t-t')^3}}\left(\nu(t')-\nu(t)\right)\,\mathrm{d}t'\right)\\ &= \nu(t)\lim_{x\to 0}\left(x\int_0^t \frac{\exp\left(-\frac{(x-\Psi(t,t'))^2}{2(t-t')}\right)}{\sqrt{2\pi(t-t')^3}}\,\mathrm{d}t'\right),\end{aligned}$$

since the second integral converges. We use the change of variables $t - t' = x^2 u$, so that

$$\lim_{x\to 0}\left(x\int_0^t \frac{\exp\left(-\frac{(x-\Psi(t,t'))^2}{2(t-t')}\right)}{\sqrt{2\pi(t-t')^3}}\,\mathrm{d}t'\right) = \lim_{x\to 0}\int_0^{t/x^2}\frac{\exp\left(-\frac{1}{2u}\right)}{\sqrt{2\pi u^3}}\,\mathrm{d}u = \int_0^\infty \frac{\exp\left(-\frac{1}{2u}\right)}{\sqrt{2\pi u^3}}\,\mathrm{d}u = 1.$$

Since the second integral converges, it is clear that

$$\mathbb{L}_1^{(2)} = \int_0^t \frac{\Psi(t,t')\,\Xi(t,t')}{\sqrt{2\pi(t-t')^3}}\nu(t')\,\mathrm{d}t'.$$

Thus, (6.2.1) is valid.

The requirement $\lim_{x\to 0} p(t,x) = 0$ thus leads to the following Volterra integral equation of the second kind for ν,

$$\nu(t) - \int_0^t \frac{\Psi(t,t')\,\Xi(t,t')}{\sqrt{2\pi(t-t')^3}}\nu(t')\,\mathrm{d}t' + \frac{\exp\left(-\frac{(M(t)+z)^2}{2t}\right)}{\sqrt{2\pi t}} = 0, \tag{6.4}$$

where

$$\begin{aligned}\Xi(t,t') &= \exp\left(-\frac{\Psi(t,t')^2}{2(t-t')}\right), \quad t \neq t',\\ \Xi(t,t) &= 1.\end{aligned}$$

6.2.2 Computation of loss rate over boundary

In this section, we compute $g(t)$ from (2.82), suppressing z, by first differenting (6.3),

$$\begin{aligned}p_x(t,x) &= \int_0^t \left(1 - \frac{(x-\Psi(t,t'))^2}{(t-t')}\right)\frac{\exp\left(-\frac{(x-\Psi(t,t'))^2}{2(t-t')}\right)}{\sqrt{2\pi(t-t')^3}}\nu(t')\,\mathrm{d}t'\\ &\quad-(x-M(t)-z)\frac{\exp\left(-\frac{(x-M(t)-z)^2}{2t}\right)}{\sqrt{2\pi t^3}}.\end{aligned}$$

We need to compute $\lim_{x\to 0} p_x(t,x)$. This limit is less standard and more difficult to compute

$$\begin{aligned}\mathbb{L}_2 := \lim_{x\to 0} p_x(t,x) &= \lim_{x\to 0}\int_0^t \left(1-\frac{\left(x-\Psi\left(t,t'\right)\right)^2}{\left(t-t'\right)}\right)\frac{\exp\left(-\frac{\left(x-\Psi(t,t')\right)^2}{2(t-t')}\right)}{\sqrt{2\pi\left(t-t'\right)^3}}\nu\left(t'\right)\,\mathrm{d}t' \\ &= 2\left(M_t\left(t\right)-\frac{1}{\sqrt{2\pi t}}\right)\nu\left(t\right)+\int_0^t \frac{\left(\left(1-\frac{\Psi(t,t')^2}{(t-t')}\right)\Xi\left(t,t'\right)\nu\left(t'\right)-\nu\left(t\right)\right)}{\sqrt{2\pi\left(t-t'\right)^3}}\,\mathrm{d}t.\end{aligned}$$

We introduce a non-singular function ϕ,

$$\phi\left(t,t'\right)=\frac{\Psi\left(t,t'\right)}{t-t'},\quad t\neq t',\qquad\qquad \phi\left(t,t\right)=M_t\left(t\right),$$

and write

$$\begin{aligned}\mathbb{L}_2 &= \lim_{x\to 0}\int_0^t \left(1-\frac{x^2}{\left(t-t'\right)}+2x\phi\left(t,t'\right)-\phi\left(t,t'\right)^2\left(t-t'\right)\right) \\ &\quad\times\frac{\exp\left(-\frac{x^2}{2(t-t')}+x\phi\left(t,t'\right)\right)\Xi\left(t,t'\right)}{\sqrt{2\pi\left(t-t'\right)^3}}\nu\left(t'\right)\,\mathrm{d}t' \\ &= \mathbb{L}_2^{(1)}+\mathbb{L}_2^{(2)}+2\mathbb{L}_2^{(3)}-\mathbb{L}_2^{(4)},\end{aligned}$$

where

$$\mathbb{L}_2^{(1)} = \lim_{x\to 0}\nu\left(t\right)\int_0^t\left(1-\frac{x^2}{\left(t-t'\right)}\right)\frac{\exp\left(-\frac{x^2}{2(t-t')}+2x\phi\left(t,t'\right)\right)}{\sqrt{2\pi\left(t-t'\right)^3}}\,\mathrm{d}t',$$

$$\mathbb{L}_2^{(2)} = \lim_{x\to 0}\int_0^t\left(1-\frac{x^2}{\left(t-t'\right)}\right)\frac{\exp\left(-\frac{x^2}{2(t-t')}+2x\phi\left(t,t'\right)\right)}{\sqrt{2\pi\left(t-t'\right)^3}}\left(\Xi\left(t,t'\right)\nu\left(t'\right)-\nu\left(t\right)\right)\,\mathrm{d}t',$$

$$\mathbb{L}_2^{(3)} = \lim_{x\to 0} x\int_0^t\phi\left(t,t'\right)\frac{\exp\left(-\frac{x^2}{2(t-t')}+2x\phi\left(t,t'\right)\right)\Xi\left(t,t'\right)}{\sqrt{2\pi\left(t-t'\right)^3}}\nu\left(t'\right)\,\mathrm{d}t',$$

$$\mathbb{L}_2^{(4)} = \lim_{x\to 0}\int_0^t\phi\left(t,t'\right)^2\frac{\exp\left(-\frac{x^2}{2(t-t')}+x\phi\left(t,t'\right)\right)\Xi\left(t,t'\right)}{\sqrt{2\pi\left(t-t'\right)}}\nu\left(t'\right)\,\mathrm{d}t'.$$

We have

$$\begin{aligned}
\mathbb{L}_2^{(1)} &= \lim_{x\to 0} \nu(t) \int_0^t \left(1 - \frac{x^2}{(t-t')}\right) \frac{\exp\left(-\frac{x^2}{2(t-t')} + 2x\phi(t,t')\right)}{\sqrt{2\pi(t-t')^3}}\,\mathrm{d}t' \\
&= \nu(t) \lim_{x\to 0} \frac{1}{x} \int_0^{t/x^2} \left(1 - \frac{1}{u}\right) \frac{\exp\left(-\frac{1}{2u}\right)}{\sqrt{2\pi u^3}}\,\mathrm{d}u \\
&= 2\nu(t) \lim_{x\to 0} \frac{1}{x} \int_{x/\sqrt{t}}^{\infty} \left(1 - v^2\right) \frac{\exp\left(-\frac{v^2}{2}\right)}{\sqrt{2\pi}}\,\mathrm{d}v \\
&= -2\nu(t) \lim_{x\to 0} \frac{1}{x} \int_0^{x/\sqrt{t}} \left(1 - v^2\right) \frac{\exp\left(-\frac{v^2}{2}\right)}{\sqrt{2\pi}}\,\mathrm{d}v \\
&= -\frac{2}{\sqrt{2\pi t}} \nu(t),
\end{aligned}$$

where $(t-t') = u$, $u = 1/v^2$ and we have used the fact that

$$\int_0^{\infty} \left(1 - v^2\right) \frac{\exp\left(-\frac{v^2}{2}\right)}{\sqrt{2\pi}}\,\mathrm{d}v = 0.$$

Further,

$$\begin{aligned}
\mathbb{L}_2^{(2)} &= \lim_{x\to 0} \int_0^t \left(1 - \frac{x^2}{(t-t')}\right) \frac{\exp\left(-\frac{x^2}{2(t-t')} + 2x\phi(t,t')\right)}{\sqrt{2\pi(t-t')^3}} \left(\Xi(t,t')\,\nu(t') - \nu(t)\right)\,\mathrm{d}t' \\
&= \lim_{x\to 0} \int_0^t \frac{\exp\left(-\frac{x^2}{2(t-t')} + 2x\phi(t,t')\right)}{\sqrt{2\pi(t-t')^3}} \left(\Xi(t,t')\,\nu(t') - \nu(t)\right)\,\mathrm{d}t' \\
&= \int_0^t \frac{\left(\Xi(t,t')\,\nu(t') - \nu(t)\right)}{\sqrt{2\pi(t-t')^3}}\,\mathrm{d}t',
\end{aligned}$$

where we have dropped the higher order x^2 term in the integral in the second line,

$$\begin{aligned}
\mathbb{L}_2^{(3)} &= \lim_{x\to 0} x \int_0^t \phi(t,t') \frac{\exp\left(-\frac{x^2}{2(t-t')} + 2x\phi(t,t')\right)}{\sqrt{2\pi(t-t')^3}} \Xi(t,t')\,\nu(t')\,\mathrm{d}t' \\
&= \phi(t,t)\,\Xi(t,t)\,\nu(t) = M_t(t)\,\nu(t),
\end{aligned}$$

and

$$\begin{aligned}
\mathbb{L}_2^{(4)} &= \lim_{x\to 0} \int_0^t \phi(t,t')^2 \frac{\exp\left(-\frac{x^2}{2(t-t')} + x\phi(t,t')\right)}{\sqrt{2\pi(t-t')}} \Xi(t,t')\,\nu(t')\,\mathrm{d}t' \\
&= \int_0^t \phi(t,t')^2 \frac{\Xi(t,t')}{\sqrt{2\pi(t-t')}} \nu(t')\,\mathrm{d}t' = \int_0^t \frac{\Psi(t,t')^2}{(t-t')} \frac{\Xi(t,t')}{\sqrt{2\pi(t-t')^3}} \nu(t')\,\mathrm{d}t'.
\end{aligned}$$

Finally,

$$\begin{aligned}\mathbb{L}_2 &= \mathbb{L}_2^{(1)} + \mathbb{L}_2^{(2)} + 2\mathbb{L}_2^{(3)} - \mathbb{L}_2^{(4)} \\ &= -2\frac{1}{\sqrt{2\pi t}}\nu(t) + \int_0^t \frac{(\Xi(t,t')\,\nu(t') - \nu(t))}{\sqrt{2\pi(t-t')^3}}\,\mathrm{d}t' \\ &+ 2M_t(t)\,\nu(t) - \int_0^t \frac{\Psi(t,t')^2}{(t-t')}\frac{\Xi(t,t')}{\sqrt{2\pi(t-t')^3}}\nu(t')\,\mathrm{d}t' \\ &= 2\left(M_t(t) - \frac{1}{\sqrt{2\pi t}}\right)\nu(t) + \int_0^t \frac{\left(\left(1-\frac{\Psi(t,t')^2}{(t-t')}\right)\Xi(t,t')\,\nu(t') - \nu(t)\right)}{\sqrt{2\pi(t-t')^3}}\,\mathrm{d}t',\end{aligned}$$

as stated.

Accordingly,

$$\begin{aligned}g(t) &= \left(M_t(t) - \frac{1}{\sqrt{2\pi t}}\right)\nu(t) \\ &+ \frac{1}{2}\int_0^t \frac{\left(\left(1-\frac{\Psi(t,t')^2}{(t-t')}\right)\Xi(t,t')\,\nu(t') - \nu(t)\right)}{\sqrt{2\pi(t-t')^3}}\mathrm{d}t' + \frac{(M(t)+z)\exp\left(-\frac{(M(t)+z)^2}{2t}\right)}{2\sqrt{2\pi t^3}}.\end{aligned} \tag{6.5}$$

6.2.3 Direct computation of loss rate

Alternatively, we can represent $g(t)$ using (2.82) by

$$g(t) = -\frac{\mathrm{d}}{\mathrm{d}t}\int_0^\infty p(t,x)\,\mathrm{d}x,$$

so that

$$\begin{aligned}g(t) &= -\frac{\mathrm{d}}{\mathrm{d}t}\int_0^t \left(\int_0^\infty \frac{(x-\Psi(t,t'))\exp\left(-\frac{(x-\Psi(t,t'))^2}{2(t-t')}\right)}{\sqrt{2\pi(t-t')^3}}\,\mathrm{d}x\right)\nu(t')\,\mathrm{d}t' \\ &- \frac{\mathrm{d}}{\mathrm{d}t}\int_0^\infty \frac{\exp\left(-\frac{(x-M(t)-z)^2}{2t}\right)}{\sqrt{2\pi t}}\,\mathrm{d}x \\ &= -\frac{\mathrm{d}}{\mathrm{d}t}\int_0^t \left(\int_{-\Psi(t,t')}^\infty \frac{\xi\exp\left(-\frac{\xi^2}{2(t-t')}\right)}{\sqrt{2\pi(t-t')^3}}\,\mathrm{d}\xi\right)\nu(t')\,\mathrm{d}t' - \frac{\mathrm{d}}{\mathrm{d}t}\int_{-\frac{(M(t)+z)}{\sqrt{t}}}^\infty \frac{\exp\left(-\frac{\xi^2}{2}\right)}{\sqrt{2\pi}}\,\mathrm{d}\xi \\ &= \frac{\mathrm{d}}{\mathrm{d}t}\int_0^t \frac{\left(\int_{-\Psi(t,t')}^\infty \mathrm{d}\left(\exp\left(-\frac{\xi^2}{2(t-t')}\right)\right)\right)}{\sqrt{2\pi(t-t')}}\nu(t')\,\mathrm{d}t' - \frac{\mathrm{d}}{\mathrm{d}t}\Phi\left(\frac{M(t)+z}{\sqrt{t}}\right) \\ &= -\frac{\mathrm{d}}{\mathrm{d}t}\int_0^t \frac{\Xi(t,t')}{\sqrt{2\pi(t-t')}}\nu(t')\,\mathrm{d}t' - \left(M_t(t) - \frac{(M(t)+z)}{2t}\right)\frac{\exp\left(-\frac{(M(t)+z)^2}{2t}\right)}{\sqrt{2\pi t}}.\end{aligned}$$

As an aside, we can verify easily by direct computation that the two expressions for g agree. We apply the following lemma, which will also be useful in Section 6.3.2.

Lemma 6.1. *Consider a differentiable function $\Xi(t,t')$ such that $\Xi(t,t)=1$. Then,*

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_0^t \frac{\Xi\left(t,t'\right)\nu\left(t'\right)}{\sqrt{2\pi\left(t-t'\right)}}\,\mathrm{d}t' = \frac{\nu\left(t\right)}{\sqrt{2\pi t}} + \frac{1}{2}\int_0^t \frac{\nu\left(t\right)-\left(\Xi\left(t,t'\right)-2\left(t-t'\right)\Xi_t\left(t,t'\right)\right)\nu\left(t'\right)}{\sqrt{2\pi\left(t-t'\right)^3}}\,\mathrm{d}t'$$
$$= \int_0^t \frac{\left(\left(\Xi\left(t,t'\right)-2\left(t-t'\right)\Xi_t\left(t,t'\right)\right)\nu\left(t'\right)\right)_{t'}}{\sqrt{2\pi\left(t-t'\right)}}\,\mathrm{d}t'.$$

Proof of Lemma 6.1. We start with the first term and judiciously use integration by parts several times to get

$$\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}t}\int_0^t \frac{\Xi\left(t,t'\right)\nu\left(t'\right)}{\sqrt{2\pi\left(t-t'\right)}}\,\mathrm{d}t' \\
&= -2\frac{\mathrm{d}}{\mathrm{d}t}\int_0^t \Xi\left(t,t'\right)\nu\left(t'\right)\mathrm{d}\sqrt{\frac{\left(t-t'\right)}{2\pi}} \\
&= 2\frac{\mathrm{d}}{\mathrm{d}t}\left(\Xi\left(t,0\right)\nu\left(0\right)\sqrt{\frac{t}{2\pi}} + \int_0^t \sqrt{\frac{\left(t-t'\right)}{2\pi}}\mathrm{d}\left(\Xi\left(t,t'\right)\nu\left(t'\right)\right)\right) \\
&= \frac{\left(2t\Xi_t\left(t,0\right)+\Xi\left(t,0\right)\right)\nu\left(0\right)}{\sqrt{2\pi t}} + \int_0^t \frac{1}{\sqrt{2\pi\left(t-t'\right)}}\mathrm{d}\left(\Xi\left(t,t'\right)\nu\left(t'\right)-\nu\left(t\right)\right) \\
&+ 2\int_0^t \sqrt{\frac{\left(t-t'\right)}{2\pi}}\mathrm{d}\left(\Xi_t\left(t,t'\right)\nu\left(t'\right)\right) \\
&= \frac{\left(2t\Xi_t\left(t,0\right)+\Xi\left(t,0\right)\right)\nu\left(0\right)}{\sqrt{2\pi t}} - \frac{\Xi\left(t,0\right)\nu\left(0\right)-\nu\left(t\right)}{\sqrt{2\pi t}} - \frac{1}{2}\int_0^t \frac{\Xi\left(t,t'\right)\nu\left(t'\right)-\nu\left(t\right)}{\sqrt{2\pi\left(t-t'\right)^3}}\,\mathrm{d}t' \\
&- 2\Xi_t\left(t,0\right)\nu\left(0\right)\sqrt{\frac{t}{2\pi}} + \int_0^t \frac{\Xi_t\left(t,t'\right)}{\sqrt{2\pi\left(t-t'\right)}}\nu\left(t'\right)\,\mathrm{d}t' \\
&= \frac{\nu\left(t\right)}{\sqrt{2\pi t}} + \frac{1}{2}\int_0^t \frac{\nu\left(t\right)-\left(\Xi\left(t,t'\right)-2\left(t-t'\right)\Xi_t\left(t,t'\right)\right)\nu\left(t'\right)}{\sqrt{2\pi\left(t-t'\right)^3}}\,\mathrm{d}t' \\
&= \frac{\nu\left(t\right)}{\sqrt{2\pi t}} + \int_0^t \left(\nu\left(t\right)-\left(\Xi\left(t,t'\right)-2\left(t-t'\right)\Xi_t\left(t,t'\right)\right)\nu\left(t'\right)\right)\mathrm{d}\left(\frac{1}{\sqrt{2\pi\left(t-t'\right)}}\right) \\
&= \int_0^t \frac{\left(\left(\Xi\left(t,t'\right)-2\left(t-t'\right)\Xi_t\left(t,t'\right)\right)\nu\left(t'\right)\right)_{t'}}{\sqrt{2\pi\left(t-t'\right)}}\,\mathrm{d}t'.
\end{aligned}$$

□

We use (6.4) and rewrite the second term in the form

$$M_t\left(t\right)\frac{\exp\left(-\frac{\left(M(t)+z\right)^2}{2t}\right)}{\sqrt{2\pi t}} = -M_t\left(t\right)\nu\left(t\right) + M_t\left(t\right)\int_0^t \frac{\Psi\left(t,t'\right)\Xi\left(t,t'\right)}{\sqrt{2\pi\left(t-t'\right)^3}}\nu\left(t'\right)\mathrm{d}t'.$$

Using the first equality in the lemma, we arrive at the following expression:

$$
\begin{aligned}
g(t) = & \left(M_t(t) - \frac{1}{\sqrt{2\pi t}}\right)\nu(t) \\
& - \frac{1}{2}\int_0^t \frac{\nu(t) - \left(\Xi(t,t') - 2\left(M_t(t)\,\Psi(t,t')\,\Xi(t,t') + (t-t')\,\Xi_t(t,t')\right)\right)\nu(t')}{\sqrt{2\pi\,(t-t')^3}}\,\mathrm{d}t' \\
& + \frac{(M(t)+z)\exp\left(-\frac{(M(t)+z)^2}{2t}\right)}{2\sqrt{2\pi t^3}}.
\end{aligned}
$$

We notice that

$$
\Xi_t(t,t') = \left(-\frac{M_t(t)\,\Psi(t,t')}{(t-t')} + \frac{\Psi(t,t')^2}{2\,(t-t')^2}\right)\Xi(t,t'),
$$

so that

$$
\Xi(t,t') - 2\left(M_t(t)\,\Psi(t,t')\,\Xi(t,t') + (t-t')\,\Xi_t(t,t')\right) = \left(1 - \frac{\Psi(t,t')^2}{(t-t')}\right)\Xi(t,t'),
$$

from which (6.5) follows.

6.3 Solution of the McKean-Vlasov equation

Now, for the McKean-Vlasov equation (2.83) we set in (6.4) and (6.5)

$$
\begin{aligned}
M(t) &= -\alpha\int_0^t g(t')\,\mathrm{d}t', \\
\Psi(t,t') &= M(t) - M(t') = -\alpha\int_{t'}^t g(t'')\,\mathrm{d}t'' = -\alpha\Omega(t,t'),
\end{aligned}
$$

to obtain a system of integral equations

$$
\left\{
\begin{aligned}
& \nu(t) + \int_0^t \frac{\alpha\Omega(t,t')\exp\left(-\frac{\alpha^2\Omega(t,t')^2}{2(t-t')}\right)}{\sqrt{2\pi\,(t-t')^3}}\nu(t')\,\mathrm{d}t' + \frac{\exp\left(-\frac{(\alpha\Omega(t,0)-z)^2}{2t}\right)}{\sqrt{2\pi t}} = 0, \\
& g(t) + \left(\alpha g(t) + \frac{1}{\sqrt{2\pi t}}\right)\nu(t) + \frac{(\alpha\Omega(t,0)-z)\exp\left(-\frac{(\alpha\Omega(t,0)-z)^2}{2t}\right)}{2\sqrt{2\pi t^3}} \\
& \qquad + \frac{1}{2}\int_0^t \frac{\left(\nu(t) - \left(1 - \frac{\alpha^2\Omega(t,t')^2}{(t-t')}\right)\exp\left(-\frac{\alpha^2\Omega(t,t')^2}{2(t-t')}\right)\nu(t')\right)}{\sqrt{2\pi\,(t-t')^3}}\,\mathrm{d}t' = 0.
\end{aligned}
\right. \tag{6.6}
$$

6.3.1 Special cases

For illustration, we work out the solution from the formula obtained in Section 6.2 for two special cases which are also accessible by the standard reflection principle for Brownian motion or method of images for parabolic equations.

6.3.1.1 $M(t) = 0$

When $M(t) = 0$, we get

$$\begin{cases} \nu(t) + \dfrac{\exp\left(-\frac{z^2}{2t}\right)}{\sqrt{2\pi t}} = 0, \\ g(t) + \dfrac{1}{\sqrt{2\pi t}}\nu(t) + \dfrac{1}{2}\displaystyle\int_0^t \dfrac{(\nu(t) - \nu(t'))}{\sqrt{2\pi(t-t')^3}}\,\mathrm{d}t' - \dfrac{z\exp\left(-\frac{z^2}{2t}\right)}{2\sqrt{2\pi t^3}} = 0. \end{cases}$$

Integration by parts of the second equation and use of the first yields

$$\begin{aligned} g(t) &= -\int_0^t \frac{\dot{\nu}(t')}{\sqrt{2\pi(t-t')}}\,\mathrm{d}t' + \frac{z\exp\left(-\frac{z^2}{2t}\right)}{2\sqrt{2\pi t^3}} \\ &= \frac{z\exp\left(-\frac{z^2}{2t}\right)}{2\sqrt{2\pi t^3}} + \frac{z\exp\left(-\frac{z^2}{2t}\right)}{2\sqrt{2\pi t^3}} \\ &= \frac{z\exp\left(-\frac{z^2}{2t}\right)}{\sqrt{2\pi t^3}}, \end{aligned}$$

as expected.

6.3.1.2 $M(t) = \mu t$

When $M(t) = \mu t$, we get

$$\begin{cases} \nu(t) - \mu\displaystyle\int_0^t \frac{\exp\left(-\frac{\mu^2(t-t')}{2}\right)}{\sqrt{2\pi(t-t')}}\nu(t')\mathrm{d}t' + \frac{\exp\left(-\frac{(\mu t+z)^2}{2t}\right)}{\sqrt{2\pi t}} = 0, \\ g(t) = \dfrac{1}{2\sqrt{2\pi}}\displaystyle\int_0^t \frac{1}{\sqrt{(t-t')^3}}\left(\exp\left(-\frac{\mu^2(t-t')}{2}\right)\nu(t') - \nu(t)\right)\mathrm{d}t' \\ -\dfrac{\mu^2}{2\sqrt{2\pi}}\displaystyle\int_0^t \frac{1}{\sqrt{t-t'}}\exp\left(-\frac{\mu^2(t-t')}{2}\right)\nu(t')\,\mathrm{d}t' + \left(\mu - \frac{1}{\sqrt{2\pi t}}\right)\nu(t) \\ +\dfrac{(\mu t+z)\exp\left(-\frac{(\mu t+z)^2}{2t}\right)}{\sqrt{2\pi t^3}}. \end{cases} \tag{6.7}$$

Taking the Laplace transform of the first equation in (6.7), we have

$$\hat{\nu}(s) - \mu\hat{\nu}(s)\frac{1}{\sqrt{2s+\mu^2}} + \frac{\exp(-z\sqrt{2s+\mu^2} - \mu z)}{\sqrt{2s+\mu^2}} = 0.$$

Hence,

$$\hat{\nu}(s) = -\frac{\exp(-z\sqrt{2s+\mu^2} - \mu z)}{\sqrt{2s+\mu^2} - \mu}.$$

The inverse Laplace transform of $\hat{\nu}(s)$ can be found analytically, but we do not need it to compute $g(t)$. Consider the second equation of (6.7). The first integral can be rewritten

as

$$\frac{1}{2\sqrt{2\pi}}\int_0^t \frac{1}{\sqrt{(t-t')^3}}\left(\exp\left(-\frac{\mu^2(t-t')}{2}\right)\nu\left(t'\right)-\nu(t)\right)\mathrm{d}t'$$
$$=\frac{1}{\sqrt{2\pi}}\int_0^t \left(\exp\left(-\frac{\mu^2(t-t')}{2}\right)\nu\left(t'\right)-\nu(t)\right)\mathrm{d}\left(\frac{1}{\sqrt{(t-t')}}\right)$$
$$=-\frac{1}{\sqrt{2\pi t}}\left(\exp\left(-\frac{\mu^2 t}{2}\right)\nu\left(0\right)-\nu(t)\right)-\frac{1}{\sqrt{2\pi}}\int_0^t \frac{1}{\sqrt{t-t'}}\exp\left(-\frac{\mu^2(t-t')}{2}\right)\dot{\nu}\left(t'\right)\mathrm{d}t'$$
$$-\frac{\mu^2}{2\sqrt{2\pi}}\int_0^t \frac{1}{\sqrt{(t-t')}}\exp\left(-\frac{\mu^2(t-t')}{2}\right)\nu\left(t'\right)\mathrm{d}t'.$$

Thus,

$$g(t)=\mu\nu(t)-\frac{1}{\sqrt{2\pi t}}\exp\left(-\frac{\mu^2 t}{2}\right)\nu\left(0\right)-\frac{1}{\sqrt{2\pi}}\int_0^t \frac{1}{\sqrt{t-t'}}\exp\left(-\frac{\mu^2(t-t')}{2}\right)\dot{\nu}\left(t'\right)\mathrm{d}t'$$
$$-\frac{\mu^2}{\sqrt{2\pi}}\int_0^t \frac{1}{\sqrt{(t-t')}}\exp\left(-\frac{\mu^2(t-t')}{2}\right)\nu\left(t'\right)\mathrm{d}t'+\frac{(\mu t+z)\exp\left(-\frac{(\mu t+z)^2}{2t}\right)}{2\sqrt{2\pi t^3}}.$$

Taking Laplace transform of the last equation, we get

$$\hat{g}(s)=\mu\hat{\nu}(s)-\frac{1}{\sqrt{2s+\mu^2}}\nu(0)-\frac{1}{\sqrt{2s+\mu^2}}\left(s\hat{\nu}(s)-\nu(0)\right)-\frac{\mu^2}{\sqrt{2s+\mu^2}}\hat{\nu}(s)$$
$$+\left(\frac{1}{2}+\frac{\mu}{2\sqrt{2s+\mu^2}}\right)\exp(-z\sqrt{2s+\mu^2}+\mu z)$$
$$=\left(-\frac{\mu}{\sqrt{2s+\mu^2}-\mu}+\frac{s+\mu^2}{\sqrt{2s+\mu^2}}\frac{1}{\sqrt{2s+\mu^2}-\mu^2}\right.$$
$$\left.+\frac{1}{2}+\frac{\mu}{2\sqrt{2s+\mu^2}}\right)\exp(-z\sqrt{2s+\mu^2}+\mu z)=\exp(-z\sqrt{2s+\mu^2}+\mu z).$$

The inverse Laplace transform yields the final expression for $g(t)$:

$$g(t)=\frac{z\exp\left(-\frac{(z-\mu t)^2}{2t}\right)}{\sqrt{2\pi t^3}},$$

as expected.

In general, only approximations to the solution can be found. We give an asymptotic and a numerical approach in the remainder of this section.

6.3.2 Perturbation solution

We expand the solution of (6.6) formally in powers of α:

$$\left(\nu\left(t\right),g\left(t\right)\right)=\left(\nu_0\left(t\right),g_0\left(t\right)\right)+\alpha\left(\nu_1\left(t\right),g_1\left(t\right)\right)+\alpha^2\left(\nu_2\left(t\right),g_2\left(t\right)\right)+...,$$

which we will truncate after the first two terms. This will give us an analytical expression which can be expected to be a good approximation for small values of α.

By plugging the formal expansion in (6.6), and collecting the terms of the same order in α, we get the following systems for $(\nu_0(t), g_0(t))$ and $(\nu_1(t), g_1(t))$:

$$\begin{cases} \nu_0(t) + \dfrac{\exp\left(-\frac{z^2}{2t}\right)}{\sqrt{2\pi t}} = 0, \\ g_0(t) + \dfrac{1}{\sqrt{2\pi t}}\nu_0(t) + \dfrac{1}{2}\displaystyle\int_0^t \frac{(\nu_0(t) - \nu_0(t'))}{\sqrt{2\pi(t-t')^3}}\mathrm{d}t' - \frac{z\exp\left(-\frac{z^2}{2t}\right)}{2\sqrt{2\pi t^3}} = 0. \end{cases}$$

$$\begin{cases} \nu_1(t) + \displaystyle\int_0^t \frac{\Omega_0(t,t')}{\sqrt{2\pi(t-t')^3}}\nu_0(t')\,\mathrm{d}t' + \frac{z\Omega_0(t,0)\exp\left(-\frac{z^2}{2t}\right)}{\sqrt{2\pi t^3}} = 0, \\ g_1(t) + g_0(t)\nu_0(t) + \dfrac{1}{\sqrt{2\pi t}}\nu_1(t) + \dfrac{1}{2}\displaystyle\int_0^t \frac{(\nu_1(t) - \nu_1(t'))}{\sqrt{2\pi(t-t')^3}}\,\mathrm{d}t' \\ \quad + \left(1 - \dfrac{z^2}{t}\right)\dfrac{\Omega_0(t,0)\exp\left(-\frac{z^2}{2t}\right)}{2\sqrt{2\pi t^3}} = 0, \end{cases}$$

where $\Omega_0(t,t') = \int_{t'}^t g_0(t'')\,\mathrm{d}t''$. The equations for g_0 and g_1 can be written as

$$g_0(t) + \int_0^t \frac{\nu_{0t}(t')}{\sqrt{2\pi(t-t')^3}}\,\mathrm{d}t' - \frac{z\exp\left(-\frac{z^2}{2t}\right)}{2\sqrt{2\pi t^3}} = 0,$$

$$g_1(t) + g_0(t)\nu_0(t) + \int_0^t \frac{\nu_{1t}(t')}{\sqrt{2\pi(t-t')}}\,\mathrm{d}t' + \left(1 - \frac{z^2}{t}\right)\frac{\Omega_0(t,0)\exp\left(-\frac{z^2}{2t}\right)}{2\sqrt{2\pi t^3}} = 0.$$

Thus, using the results in Section 6.3.1 for $\alpha = 0$,

$$\nu_0(t) = -\frac{\exp\left(-\frac{z^2}{2t}\right)}{\sqrt{2\pi t}}, \qquad g_0(t) = \frac{z\exp\left(-\frac{z^2}{2t}\right)}{\sqrt{2\pi t^3}}$$
$$\Omega_0(t,t') = 2\left(\Phi\left(\frac{z}{\sqrt{t'}}\right) - \Phi\left(\frac{z}{\sqrt{t}}\right)\right),$$
$$\Omega_{0t}(t,t') = \frac{z\exp\left(-\frac{z^2}{2t}\right)}{\sqrt{2\pi t^3}} = g_0(t), \qquad \Omega_{0t'}(t,t') = -g_0(t').$$

Next,

$$\begin{aligned} \nu_1(t) &= -\int_0^t \frac{\Omega_0(t,t')}{\sqrt{2\pi(t-t')^3}}\nu_0(t')\,\mathrm{d}t' - \frac{z\Omega_0(t,0)\exp\left(-\frac{z^2}{2t}\right)}{\sqrt{2\pi t^3}} \\ &= 2\int_0^t \frac{\left(\Phi\left(\frac{z}{\sqrt{t}}\right) - \Phi\left(\frac{z}{\sqrt{t'}}\right)\right)}{\sqrt{2\pi(t-t')^3}}\frac{\exp\left(-\frac{z^2}{2t'}\right)}{\sqrt{2\pi t'}}\,\mathrm{d}t' + \frac{2z\left(\Phi\left(\frac{z}{\sqrt{t}}\right) - 1\right)\exp\left(-\frac{z^2}{2t}\right)}{\sqrt{2\pi t^3}}, \end{aligned} \tag{6.8}$$

$$g_1(t) = -g_0(t)\nu_0(t) - \int_0^t \frac{\nu_{1t}(t')}{\sqrt{2\pi(t-t')}}\,\mathrm{d}t' - \left(1 - \frac{z^2}{t}\right)\frac{\Omega_0(t,0)\exp\left(-\frac{z^2}{2t}\right)}{2\sqrt{2\pi t^3}},$$

where

$$\Omega_0(t,0) = 2\left(1 - \Phi\left(\frac{z}{\sqrt{t}}\right)\right).$$

We can write $\nu_1(t)$ in the form

$$\nu_1(t) = -\int_0^t \frac{\omega_0(t,t')}{\sqrt{2\pi(t-t')}} g_0(t')\,\nu_0(t')\,\mathrm{d}t' - \frac{z\Omega_0(t,0)\exp\left(-\frac{z^2}{2t}\right)}{\sqrt{2\pi t^3}},$$

where

$$\omega_0(t,t') = \frac{\Omega_0(t,t')}{g_0(t')(t-t')} = -\frac{2\left(\Phi\left(\frac{z}{\sqrt{t}}\right) - \Phi\left(\frac{z}{\sqrt{t'}}\right)\right)\sqrt{2\pi t'^3}\exp\left(\frac{z^2}{2t'}\right)}{z(t-t')}, \quad t \neq t',$$
$$\omega_0(t,t) = 1,$$

and obtain an expression for ν_{1t}:

$$\nu_{1t}(t) = -\frac{\mathrm{d}}{\mathrm{d}t}\int_0^t \frac{\omega_0(t,t')}{\sqrt{2\pi(t-t')}} g_0(t')\nu_0(t')\,\mathrm{d}t' - \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{z\Omega_0(t,0)\exp\left(-\frac{z^2}{2t}\right)}{\sqrt{2\pi t^3}}\right). \tag{6.9}$$

To compute the first term of (6.9), we use the second equality in Lemma 6.1 with $\nu(t) = \nu_0(t)g_0(t)$ and $\Xi(t,t') = \omega_0(t,t')$, to get

$$\begin{aligned}
\nu_{1t}(t) &= -\int_0^t \frac{\left(\left(\omega_0(t,t') - 2(t-t')\,\omega_{0t}(t,t')\right) g_0(t')\,\nu_0(t')\right)_{t'}}{\sqrt{2\pi(t-t')}}\,\mathrm{d}t' \\
&\quad - \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{z\Omega_0(t,0)\exp\left(-\frac{z^2}{2t}\right)}{\sqrt{2\pi t^3}}\right) \\
&= -\int_0^t \frac{\left(\left(\frac{3\Omega_0(t,t')}{(t-t')} - 2\Omega_{0t}(t,t')\right)\nu_0(t')\right)_{t'}}{\sqrt{2\pi(t-t')}}\,\mathrm{d}t' - \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{z\Omega_0(t,0)\exp\left(-\frac{z^2}{2t}\right)}{\sqrt{2\pi t^3}}\right) \\
&= \int_0^t \frac{\left(\frac{3g_0(t')}{(t-t')} - 3\frac{\Omega_0(t,t')}{(t-t')^2}\right)\nu_0(t') - \left(3\frac{\Omega_0(t,t')}{(t-t')} - 2g_0(t)\right)\nu_{0t}(t')}{\sqrt{2\pi(t-t')}}\,\mathrm{d}t' \\
&\quad - \frac{z^2\exp\left(-\frac{z^2}{t}\right)}{2\pi t^3} + 2z\left(1 - \Phi\left(\frac{z}{\sqrt{t}}\right)\right)\left(3 - \frac{z^2}{t}\right)\frac{\exp\left(-\frac{z^2}{2t}\right)}{2\sqrt{2\pi t^5}}.
\end{aligned} \tag{6.10}$$

Substituting this in the second equation (6.8) yields an expression for $g_1(t)$, which can be evaluated by numerical integration.

Finally, we evaluate the complexity of the computation of $g_1(t)$. Consider numerical quadrature with N points. First, we precompute $\nu_{1t}(t)$ using (6.10); it can be done in $O(N^2)$ operations. Then, we can compute $g_1(t)$ using the second equation in (6.8) with precomputed $\nu_{1t}(t)$ again in $O(N^2)$. Thus, the total complexity of the perturbation method is $O(N^2)$.

6.3.3 Numerical solution

In this section, we present (without convergence analysis) a simple method for the numerical approximation of the solution to the coupled Volterra equations (6.6). We note that Volterra equations and their numerical solution are a well-established research field. For a relevant discussion of the stability and convergence of some methods for equations with a weak singularity see Linz (1985). Noble (1969) discusses possible instabilities of multi-step methods in the presence of weak singularities.

A number of papers propose higher order methods and collocation techniques to improve the convergence and treat instabilities. For example, Brunner (1985) proved the convergence for a polynomial spline collocation method with quadratures; Kolk et al. (2009), Kolk and Pedas (2009), and Kolk and Pedas (2013) used a piecewise polynomial collocation method to solve a Volterra equation with weak singularity, and derived optimal global convergence estimates and a local superconvergence result. An alternative is to consider a special functional basis, such as Chebyshev polynomials and Bernstein polynomials (see Maleknejad et al. (2007) and Maleknejad et al. (2011), respectively). In both cases, the approximation leads to a system of linear or nonlinear algebraic equations. Hairer et al. (1985) developed a method based on fast Fourier transform to reduce the number of kernel evaluations on an N-point grid from $O(N^2)$ to $O(N(\log N)^2)$.

In this paper, for simplicity, we consider trapezoidal quadrature, with a special treatment of the interval containing the singularity, to obtain the numerical solution recursively. We divide the interval $[0,T]$ into equally spaced subintervals of length Δ and discretize (6.6) appropriately. To this end, we assume that ν and g are piecewise linear with $\nu(l\Delta)=\nu_l$ and $g(l\Delta)=g_l$, so that on the interval $[(l-1)\,\Delta, l\Delta]$ we have

$$\nu(t)=\frac{\nu_{l-1}\,(l\Delta-t)+\nu_l\,(t-(l-1)\,\Delta)}{\Delta}=\nu_l-\frac{(\nu_l-\nu_{l-1})}{\Delta}\,(l\Delta-t)\,,$$
$$g(t)=\frac{g_{l-1}\,(l\Delta-t)+g_l\,(t-(l-1)\,\Delta)}{\Delta}=g_l-\frac{(g_l-g_{l-1})}{\Delta}\,(l\Delta-t)\,.$$

Accordingly,

$$\Omega_{nl}=\int_{l\Delta}^{n\Delta} g(t'')\,\mathrm{d}t'=\frac{\Delta}{2}\,(g_l+2g_{l+1}+...+2g_{n-1}+g_n)\,,$$
$$\Omega_{nl}=\Omega_{(n-1)l}+\frac{\Delta}{2}\,(g_n+g_{n-1})\,.$$

Inserting in (6.6), the discretized system of equations has the form

$$\begin{cases}\nu_n+\alpha\displaystyle\sum_{l=1}^{n}\mathcal{I}_l^n+\frac{\exp\left(-\frac{(\alpha\Omega_{n0}-z)^2}{2n\Delta}\right)}{\sqrt{2\pi n\Delta}}=0,\\[2ex] g_n+\left(\alpha g_n+\dfrac{1}{\sqrt{2\pi n\Delta}}\right)\nu_n+\dfrac{1}{2}\displaystyle\sum_{l=1}^{n}\mathcal{J}_l^n+\dfrac{(\alpha\Omega_{n0}-z)\exp\left(-\frac{(\alpha\Omega_{n0}-z)^2}{2n\Delta}\right)}{2\sqrt{2\pi n^3\Delta^3}}=0.\end{cases}\tag{6.11}$$

For a given n, all the relevant integrals $\mathcal{I}_l, \mathcal{J}_l, 1 \leq l < n$, can be approximated by the trapezoidal rule (or via more accurate composite formulas, if necessary). Accordingly,

$$\begin{aligned}\mathcal{I}_l^n &= \int_{(l-1)\Delta}^{l\Delta} \frac{\Omega\left(n\Delta, t'\right) \exp\left(-\frac{\alpha^2 \Omega(n\Delta,t')^2}{2(n\Delta-t')}\right) \nu\left(t'\right)}{\sqrt{2\pi\left(n\Delta-t'\right)^3}} \mathrm{d}t' \\ &\approx \frac{1}{\sqrt{8\pi\Delta}} \left(\frac{\Omega_{nl} \exp\left(-\frac{\alpha^2 \Omega_{nl}^2}{2(n-l)\Delta}\right) \nu_l}{(n-l)^{3/2}} + \frac{\Omega_{n(l-1)} \exp\left(-\frac{\alpha^2 \Omega_{n(l-1)}^2}{2(n-l+1)\Delta}\right) \nu_{l-1}}{(n-l+1)^{3/2}} \right),\end{aligned} \tag{6.12}$$

and

$$\begin{aligned}\mathcal{J}_l^n &= \int_{(l-1)\Delta}^{l\Delta} \frac{\left(\nu_n - \left(1 - \frac{\alpha^2 \Omega(n\Delta,t')^2}{(n\Delta-t')}\right) \exp\left(-\frac{\alpha^2 \Omega(n\Delta,t')^2}{2(n\Delta-t')}\right) \nu\left(t'\right)\right)}{\sqrt{2\pi\left(n\Delta-t'\right)^3}} \mathrm{d}t' \\ &\approx \frac{1}{\sqrt{8\pi\Delta}} \left(\frac{\left(\nu_n - \left(1 - \frac{\alpha^2 \Omega_{nl}^2}{(n-l)\Delta}\right) \exp\left(-\frac{\alpha^2 \Omega_{nl}^2}{2(n-l)\Delta}\right) \nu_l\right)}{(n-l)^{3/2}} \right. \\ &\left. + \frac{\left(\nu_n - \left(1 - \frac{\alpha^2 \Omega_{n(l-1)}^2}{(n-l+1)\Delta}\right) \exp\left(-\frac{\alpha^2 \Omega_{n(l-1)}^2}{2(n-l+1)\Delta}\right) \nu_{l-1}\right)}{(n-l+1)^{3/2}} \right).\end{aligned} \tag{6.13}$$

However, the last two integrals, $\mathcal{I}_n^n, \mathcal{J}_n^n$, require special care, because they have weak singularities. Consider the integral $\mathcal{I}_n$, which has the form

$$\mathcal{I}_n^n = \int_{(n-1)\Delta}^{n\Delta} \frac{\Omega\left(n\Delta, t'\right) \exp\left(-\frac{\alpha^2 \Omega(n\Delta,t')^2}{2(n\Delta-t')}\right)}{\sqrt{2\pi\left(n\Delta-t'\right)^3}} \nu\left(t'\right) \mathrm{d}t'.$$

In view of our piecewise linearity assumption, we have

$$\Omega\left(n\Delta, t'\right) = g_n \tau - \frac{g_n - g_{n-1}}{2\Delta} \tau^2, \tag{6.14}$$

where $\tau = n\Delta - t'$. Accordingly,

$$\mathcal{I}_n^n = \int_0^{\Delta} \frac{\left(g_n - \frac{g_n - g_{n-1}}{2\Delta}\tau\right) \exp\left(-\frac{\alpha^2 \left(g_n - \frac{(g_n - g_{n-1})}{2\Delta}\tau\right)^2 \tau}{2}\right) \left(\nu_n - \frac{(\nu_n - \nu_{n-1})}{\Delta}\tau\right)}{\sqrt{2\pi\tau}} \mathrm{d}\tau.$$

A standard change of variables $\tau = u^2$ yields

$$\begin{aligned}\mathcal{I}_n^n = \frac{2}{\sqrt{2\pi}} \int_0^{\sqrt{\Delta}} \left(g_n - \frac{(g_n - g_{n-1})}{2\Delta} u^2\right) \exp\left(-\frac{\alpha^2 \left(g_n - \frac{(g_n - g_{n-1})}{2\Delta} u^2\right)^2 u^2}{2}\right) \\ \times \left(\nu_n - \frac{(\nu_n - \nu_{n-1})}{\Delta} u^2\right) \mathrm{d}u.\end{aligned} \tag{6.15}$$

The latter integral is now non-singular and can be approximated by the trapezoidal rule:

$$\mathcal{I}_n^n \approx \frac{1}{\sqrt{2\pi\Delta}}\left(\Delta g_n \nu_n + \Gamma_n \exp\left(-\frac{\alpha^2\Gamma_n^2}{2\Delta}\right)\nu_{n-1}\right) = \sqrt{\frac{\Delta}{2\pi}} g_n \nu_n + \frac{\Gamma_n \exp\left(-\frac{\alpha^2\Gamma_n^2}{2\Delta}\right)}{\sqrt{2\pi\Delta}}\nu_{n-1}, \tag{6.16}$$

where

$$\Gamma_n = \frac{\Delta}{2}\left(g_n + g_{n-1}\right).$$

Similarly,

$$\begin{aligned}
\mathcal{J}_n^n &= \int_{(n-1)\Delta}^{n\Delta} \frac{\left(\nu_n - \left(1 - \alpha^2 \frac{\Omega(n\Delta,t')^2}{(n\Delta - t')}\right)\exp\left(-\frac{\alpha^2\Omega(n\Delta,t')^2}{2(n\Delta-t')}\right)\nu\left(t'\right)\right)}{\sqrt{2\pi\left(n\Delta - t'\right)^3}}\,\mathrm{d}t' \\
&= \int_{(n-1)\Delta}^{n\Delta} \frac{\left(\nu_n - \exp\left(-\frac{\alpha^2\Omega(n\Delta,t')^2}{2(n\Delta-t')}\right)\nu\left(t'\right)\right)}{\sqrt{2\pi\left(n\Delta - t'\right)^3}}\,\mathrm{d}t' \\
&+ \alpha^2 \int_{(n-1)\Delta}^{n\Delta} \frac{\Omega\left(n\Delta, t'\right)^2 \exp\left(-\frac{\alpha^2\Omega(n\Delta,t')^2}{2(n\Delta-t')}\right)\nu\left(t'\right)}{\left(n\Delta - t'\right)\sqrt{2\pi\left(n\Delta - t'\right)^3}}\,\mathrm{d}t' \\
&= \mathcal{J}_n^{n,1} + \alpha^2 \mathcal{J}_n^{n,2}.
\end{aligned}$$

We have

$$\begin{aligned}
\mathcal{J}_n^{n,1} &= \int_0^\Delta \frac{\left(\nu_n - \exp\left(-\frac{\alpha^2\left(g_n - \frac{(g_n - g_{n-1})}{2\Delta}\tau\right)^2\tau}{2}\right)\left(\nu_n - \frac{(\nu_n - \nu_{n-1})}{\Delta}\tau\right)\right)}{\sqrt{2\pi\tau^3}}\,\mathrm{d}\tau \\
&= \int_0^\Delta \frac{\left(\left(\left(1 - \exp\left(-\frac{\alpha^2\left(g_n - \frac{(g_n - g_{n-1})}{2\Delta}\tau\right)^2\tau}{2}\right)\right)\right)\nu_n\right)}{\sqrt{2\pi\tau^3}}\,\mathrm{d}\tau \\
&+ \frac{(\nu_n - \nu_{n-1})}{\Delta}\int_0^\Delta \frac{\exp\left(-\frac{\alpha^2\left(g_n - \frac{(g_n - g_{n-1})}{2\Delta}\tau\right)^2\tau}{2}\right)}{\sqrt{2\pi\tau}}\,\mathrm{d}\tau \\
&\approx \frac{\alpha^2\nu_n}{2}\int_0^\Delta \frac{\left(g_n - \frac{(g_n - g_{n-1})}{2\Delta}\tau\right)^2}{\sqrt{2\pi\tau}}\,\mathrm{d}\tau \qquad (6.17) \\
&+ \frac{(\nu_n - \nu_{n-1})}{\Delta}\int_0^\Delta \frac{\exp\left(-\frac{\alpha^2\left(g_n - \frac{(g_n - g_{n-1})}{2\Delta}\tau\right)^2\tau}{2}\right)}{\sqrt{2\pi\tau}}\,\mathrm{d}\tau \\
&= \frac{\alpha^2\nu_n}{\sqrt{2\pi}}\int_0^{\sqrt{\Delta}}\left(g_n - \frac{(g_n - g_{n-1})}{2\Delta}u^2\right)^2\mathrm{d}u \\
&+ \frac{2\left(\nu_n - \nu_{n-1}\right)}{\sqrt{2\pi}\Delta}\int_0^{\sqrt{\Delta}}\exp\left(-\frac{\alpha^2\left(g_n - \frac{(g_n - g_{n-1})}{2\Delta}u^2\right)^2u^2}{2}\right)\mathrm{d}u \\
&\approx \frac{\alpha^2}{2}\frac{1}{\sqrt{2\pi\Delta^3}}\left(\Delta^2 g_n^2 + \Gamma_n^2\right)\nu_n + \frac{1}{\sqrt{2\pi\Delta}}\left(1 + \exp\left(-\frac{\alpha^2\Gamma_n^2}{2\Delta}\right)\right)\left(\nu_n - \nu_{n-1}\right), \\
\mathcal{J}_n^{n,2} &= \int_0^\Delta \frac{\left(g_n - \frac{g_n - g_{n-1}}{2\Delta}\tau\right)^2\exp\left(-\frac{\alpha^2\left(g_n - \frac{(g_n - g_{n-1})}{2\Delta}\tau\right)^2\tau}{2}\right)\left(\nu_n - \frac{(\nu_n - \nu_{n-1})}{\Delta}\tau\right)}{\sqrt{2\pi\tau}}\,\mathrm{d}\tau \\
&= \frac{2}{\sqrt{2\pi}}\int_0^{\sqrt{\Delta}}\left(g_n - \frac{(g_n - g_{n-1})}{2\Delta}u^2\right)^2\exp\left(-\frac{\alpha^2\left(g_n - \frac{(g_n - g_{n-1})}{2\Delta}u^2\right)^2u^2}{2}\right) \qquad (6.18) \\
&\times \left(\nu_n - \frac{(\nu_n - \nu_{n-1})}{\Delta}u^2\right)\mathrm{d}u \\
&\approx \frac{1}{\sqrt{2\pi\Delta^3}}\left(\Delta^2 g_n^2\nu_n + \Gamma_n^2\exp\left(-\frac{\alpha^2\Gamma_n^2}{2\Delta}\right)\nu_{n-1}\right).
\end{aligned}$$

Thus,

$$\begin{aligned}\mathcal{J}_n^n &\approx \frac{1}{\sqrt{2\pi\Delta}}\left(1+\exp\left(-\frac{\alpha^2\Gamma_n^2}{2\Delta}\right)\right)(\nu_n-\nu_{n-1})\\ &+\frac{\alpha^2}{\sqrt{2\pi\Delta^3}}\left(\left(\frac{3}{2}\Delta^2 g_n^2+\Gamma_n^2\right)\nu_n+\Gamma_n^2\exp\left(-\frac{\alpha^2\Gamma_n^2}{2\Delta}\right)\nu_{n-1}\right)\\ &=\left(\frac{\left(1+\exp\left(-\frac{\alpha^2\Gamma_n^2}{2\Delta}\right)\right)}{\sqrt{2\pi\Delta}}+\frac{\alpha^2\left(\frac{3}{2}\Delta^2 g_n^2+\Gamma_n^2\right)}{\sqrt{2\pi\Delta^3}}\right)\nu_n\\ &-\left(\frac{\left(1+\exp\left(-\frac{\alpha^2\Gamma_n^2}{2\Delta}\right)\right)}{\sqrt{2\pi\Delta}}-\frac{\alpha^2\Gamma_n^2\exp\left(-\frac{\alpha^2\Gamma_n^2}{2\Delta}\right)}{\sqrt{2\pi\Delta^3}}\right)\nu_{n-1}.\end{aligned} \tag{6.19}$$

By using (6.14) we can represent expressions (6.12), (6.13) in a recurrent form, neglecting now quadrature errors:

$$\begin{aligned}\mathcal{I}_l^n &= \frac{1}{\sqrt{8\pi\Delta}}\left(\frac{\Omega_{nl}\exp\left(-\frac{\alpha^2\Omega_{nl}^2}{2(n-l)\Delta}\right)\nu_l}{(n-l)^{3/2}}+\frac{\Omega_{n(l-1)}\exp\left(-\frac{\alpha^2\Omega_{n(l-1)}^2}{2(n-l+1)\Delta}\right)\nu_{l-1}}{(n-l+1)^{3/2}}\right)\\ &=\mathbb{I}_l^n\left(g_n|\,\nu_{l-1},\nu_l,g_1,...,g_{n-1}\right),\end{aligned}$$

$$\begin{aligned}\mathcal{J}_l^n &= \frac{1}{\sqrt{8\pi\Delta}}\left(\frac{\left(\nu_n-\left(1-\frac{\alpha^2\Omega_{nl}^2}{(n-l)\Delta}\right)\exp\left(-\frac{\alpha^2\Omega_{nl}^2}{2(n-l)\Delta}\right)\nu_l\right)}{(n-l)^{3/2}}\right.\\ &\left.+\frac{\left(\nu_n-\left(1-\frac{\alpha^2\Omega_{n(l-1)}^2}{(n-l+1)\Delta}\right)\exp\left(-\frac{\alpha^2\Omega_{n(l-1)}^2}{2(n-l+1)\Delta}\right)\nu_{l-1}\right)}{(n-l+1)^{3/2}}\right)\\ &=\left(\frac{1}{\sqrt{8\pi\Delta}\,(n-l)^{3/2}}+\frac{1}{\sqrt{8\pi\Delta}\,(n-l+1)^{3/2}}\right)\nu_n\\ &-\frac{\left(1-\frac{\alpha^2\Omega_{nl}^2}{(n-l)\Delta}\right)\exp\left(-\frac{\alpha^2\Omega_{nl}^2}{2(n-l)\Delta}\right)\nu_l}{\sqrt{8\pi\Delta}\,(n-l)^{3/2}}-\frac{\left(1-\frac{\alpha^2\Omega_{n(l-1)}^2}{(n-l+1)\Delta}\right)\exp\left(-\frac{\alpha^2\Omega_{n(l-1)}^2}{2(n-l+1)\Delta}\right)\nu_{l-1}}{\sqrt{8\pi\Delta}\,(n-l+1)^{3/2}}\\ &=\mathbb{J}_l^n\left(\nu_n,g_n|\,\nu_{l-1},\nu_l,g_1,...,g_{n-1}\right)=\mathbb{U}_l^n\nu_n+\mathbb{V}_l^n\left(g_n|\,\nu_{l-1},\nu_l,g_1,...,g_{n-1}\right).\end{aligned}$$

By the same token, $\mathcal{I}_n^n$, $\mathcal{J}_n^n$ given by (6.16), (6.19) can be written in the form

$$\begin{aligned}\mathcal{I}_n^n &= \mathbb{I}_n^n\left(\nu_n,g_n|\,\nu_{n-1},g_{n-1}\right)\\ &=\mathbb{A}_n^n\left(g_n|\,g_{n-1}\right)\nu_n+\mathbb{B}_n^n\left(g_n|\,\nu_{n-1},g_{n-1}\right),\end{aligned}$$

$$\begin{aligned}\mathcal{J}_n^n &= \mathbb{J}_n^n\left(\nu_n,g_n|\,\nu_{n-1},g_{n-1}\right)\\ &=\mathbb{U}_n^n\left(g_n|\,g_{n-1}\right)\nu_n+\mathbb{V}_n^n\left(g_n|\,\nu_{n-1},g_{n-1}\right).\end{aligned}$$

In view of the above, system (6.11) can be written as follows:

$$\begin{cases} \nu_n + \alpha \sum_{l=1}^{n-1} \mathbb{I}_l^n (g_n) + \alpha \mathbb{I}_n^n (\nu_n, g_n) + \dfrac{\exp\left(-\frac{(\alpha\Omega_{n0}-z)^2}{2n\Delta}\right)}{\sqrt{2\pi n\Delta}} = 0, \\ g_n + \left(\alpha g_n + \dfrac{1}{\sqrt{2\pi n\Delta}}\right) \nu_n + \dfrac{1}{2} \sum_{l=1}^{n-1} \mathbb{J}_l^n (\nu_n, g_n) + \dfrac{1}{2} \mathbb{J}_n^n (\nu_n, g_n) \\ + \dfrac{(\alpha\Omega_{n0} - z) \exp\left(-\frac{(\alpha\Omega_{n0}-z)^2}{2n\Delta}\right)}{2\sqrt{2\pi n^3\Delta^3}} = 0, \end{cases} \tag{6.20}$$

where we suppress explicit dependencies on $g_1, ..., g_{n-1}, \nu_1, ..., \nu_{n-1}$ for brevity. provided that $g_1, ..., g_{n-1}, \nu_1, ..., \nu_{n-1}$ are given. This system can be solved by using the Newton-Raphson method, say. As a result, the new pair (ν_n, g_n) can be found and the recurrence advanced by one more step as required.

If so desired, system (6.20) can be simplified further. We notice that the dependence on ν_n is linear, eliminate ν_n in favor of g_n from the first equation,

$$\nu_n = -\frac{\alpha \left(\sum_{l=1}^{n-1} \mathbb{I}_l^n (g_n) + \mathbb{B}_n^n (g_n) \right) + \frac{\exp\left(-\frac{(\alpha\Omega_{n0}-z)^2}{2n\Delta}\right)}{\sqrt{2\pi n\Delta}}}{(1 + \alpha \mathbb{A}_n^n (g_n))},$$

and substitute this expression in the second equation, obtaining a scalar recursive non-linear equation of the form

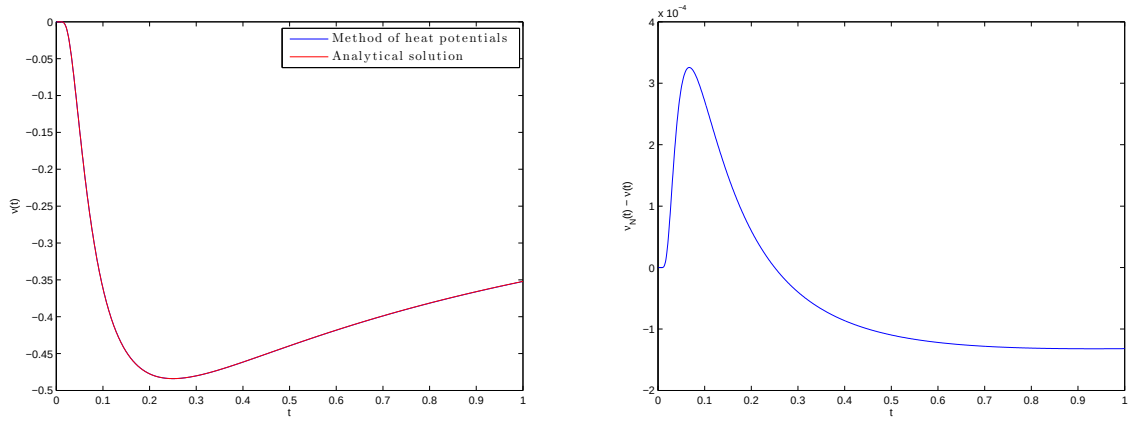
$$\begin{aligned} & g_n + \frac{1}{2} \left(\sum_{l=1}^{n-1} \mathbb{V}_l^n (g_n) + \mathbb{V}_n^n (g_n) \right) - \frac{\left(\alpha g_n + \frac{1}{\sqrt{2\pi n\Delta}} + \frac{1}{2} \left(\sum_{l=1}^{n-1} \mathbb{U}_l^n + \mathbb{U}_n^n (g_n) \right) \right)}{(1 + \alpha \mathbb{A}_n^n (g_n))} \\ & \times \left(\alpha \left(\sum_{l=1}^{n-1} \mathbb{I}_l^n (g_n) + \mathbb{B}_n^n (g_n) \right) + \frac{\exp\left(-\frac{(\alpha\Omega_{n0}-z)^2}{2n\Delta}\right)}{\sqrt{2\pi n\Delta}} \right) \\ & + \frac{(\alpha\Omega_{n0} - z) \exp\left(-\frac{(\alpha\Omega_{n0}-z)^2}{2n\Delta}\right)}{2\sqrt{2\pi n^3\Delta^3}} = 0. \end{aligned} \tag{6.21}$$

6.4 Numerical tests and results

In this section, we first analyse the convergence (order) of the numerical method, then test the accuracy of the first order expansion against the numerical solution, and finally perform parameter studies (in α) to investigate the influence of the mean-field interaction on the behaviour of the solution.

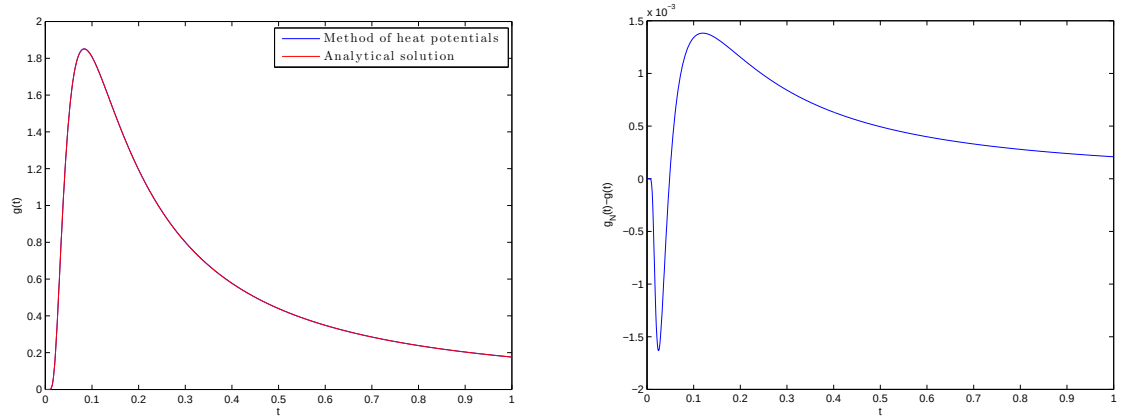
6.4.1 Numerical method

To demonstrate the performance of the discretisation scheme, we compare the solution with (6.3.1.1), the analytical solution, in the case $\alpha = 0$.



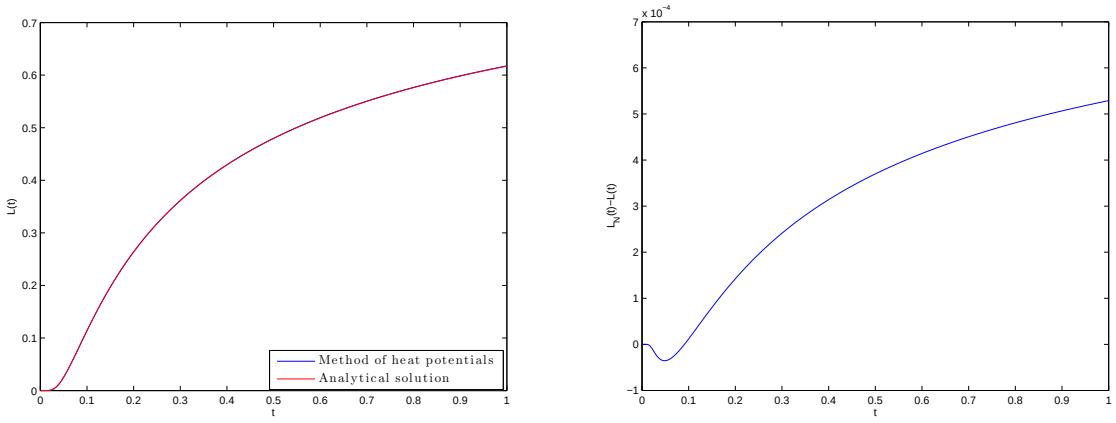
(a) (b)

Figure 6.1: $\nu(t)$: (a) Numerical and analytical solution for $\alpha = 0$, $N = 1000$ (visually indistinguishable). (b) The difference with the exact solution for $\alpha = 0$.



(a) (b)

Figure 6.2: $g(t)$: (a) Numerical and analytical solution for $\alpha = 0$, $N = 1000$ (visually indistinguishable). (b) The difference with the exact solution for $\alpha = 0$.

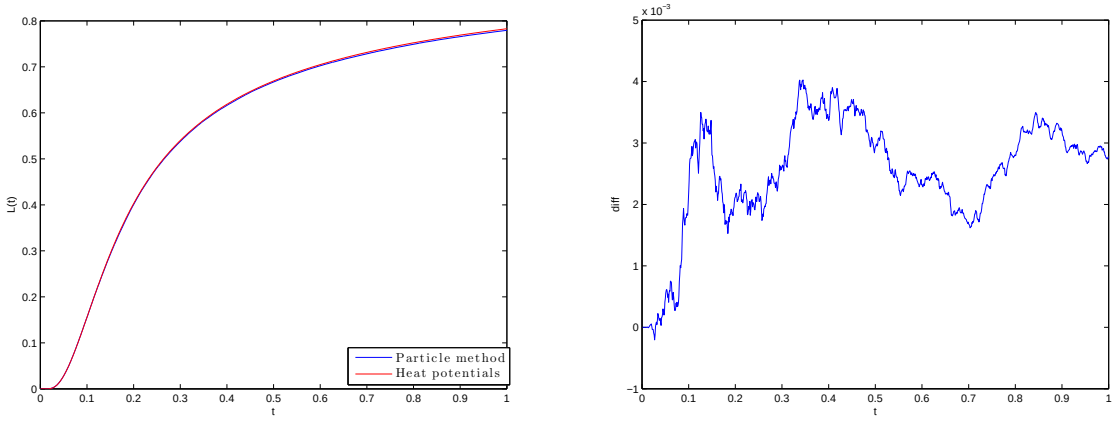


(a) (b)

Figure 6.3: L_t: (a) Numerical and analytical solution for $\alpha = 0$, $N = 1000$ (visually indistinguishable). (b) The difference with the exact solution for $\alpha = 0$.

We use the method described in Chapter 5 with sufficiently many particles and timesteps as a benchmark and compare the results.

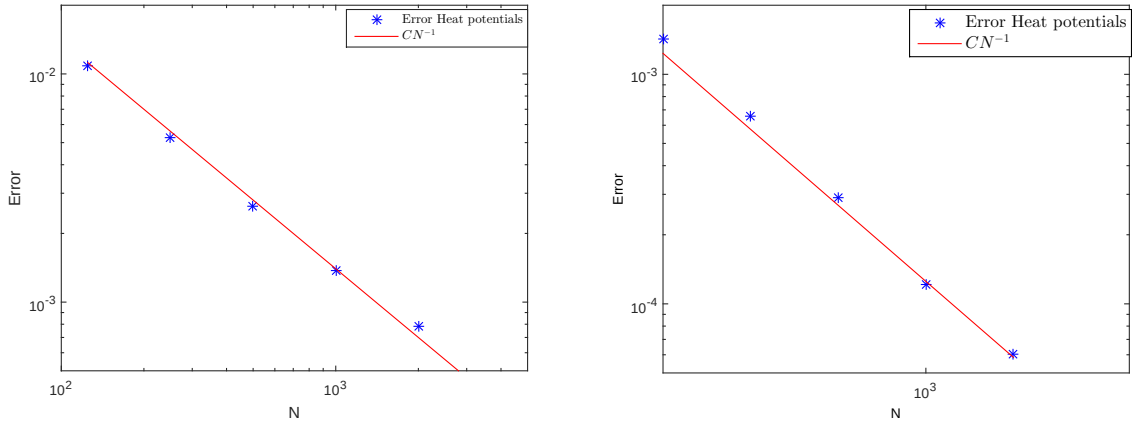
We illustrate the difference between the method described in this chapter and Chapter 5 in Figure 6.4.



(a) (b)

Figure 6.4: L_t: Comparison of the numerical solution by the method of heat potentials ($N = 1000$) with that of the particle method described in Chapter 5 for $\alpha = 0.5$: (a) The solutions are visually indistinguishable (b) The difference between the solutions.

We now analyse the convergence order of the discretisation scheme for the Volterra equation empirically. With N time steps, the error of the approximation (6.21) is expected to be $O(N^{-1})$, because the trapezoidal integration in (6.15), (6.17), and (6.18) is on intervals $(0, \sqrt{\Delta})$, and the result is divided by $\sqrt{\Delta}$ after that. We empirically confirm this in Figure 6.5.



(a) (b)

Figure 6.5: Error in the maximum norm for numerical method in Section 6.3.3: (a) for g compared to the exact solution for $\alpha = 0$; (b) for L compared to the simulation solution for $\alpha = 0.5$.

The complexity of our method is $O(N^2)$. Hence, in order to achieve precision ε, we need $O(\varepsilon^{-2})$ operations. In comparison, the particle method with Brownian bridge in Chapter 5 requires $O(\varepsilon^{-3})$ operations. The latter could be improved to $O(\varepsilon^{-2})$ by multi-level simulation, but equally a higher order method for the Volterra equation would bring the complexity down. Another advantage of the method above is that we automatically get directly the derivative g of the loss function, which is harder to obtain in simulation method because of Monte Carlo noise.

6.4.2 Comparison of perturbation and numerical methods

Here, we compare the numerical and perturbation solutions described above. We fix $T = 1$, $z = 0.5$, and choose $N = 1000$, the number of grid points, sufficiently large so that the numerical error is negligible. In Figure 6.6 we plot g, the hitting time density, computed with numerical and perturbation methods as well as $g_0(t)$, the solution for $\alpha = 0$, to measure the impact of the nonlinear term, for different values of α. For $\alpha = 0.1$, the two solutions are visually indistinguishable; for $\alpha = 0.3$ there is small but visible difference between the solutions, which increases further for $\alpha = 0.5$ and arises from the higher order terms.[1] For $\alpha = 1$, where the numerical solution shows a jump in the loss function at around $t = 0.1$, the first order expansion approximation breaks down.

6.4.3 Parameter studies

We now assess the impact of the mean-field interaction by varying the parameter α.

We fix $T = 1, z = 0.5$, and choose $N = 1000$, the number of the grid points. Figure 6.7 demonstrates the behavior of $\nu(t)$ and $g(t)$ for different values of α, starting with $\alpha = 0$; for L_t, including a case with discontinuity.

[1] In our implementation of the perturbation solution, we also perform a scaling to ensure the correct cumulative density at T (which in practice is unknown) to improve the results slightly.

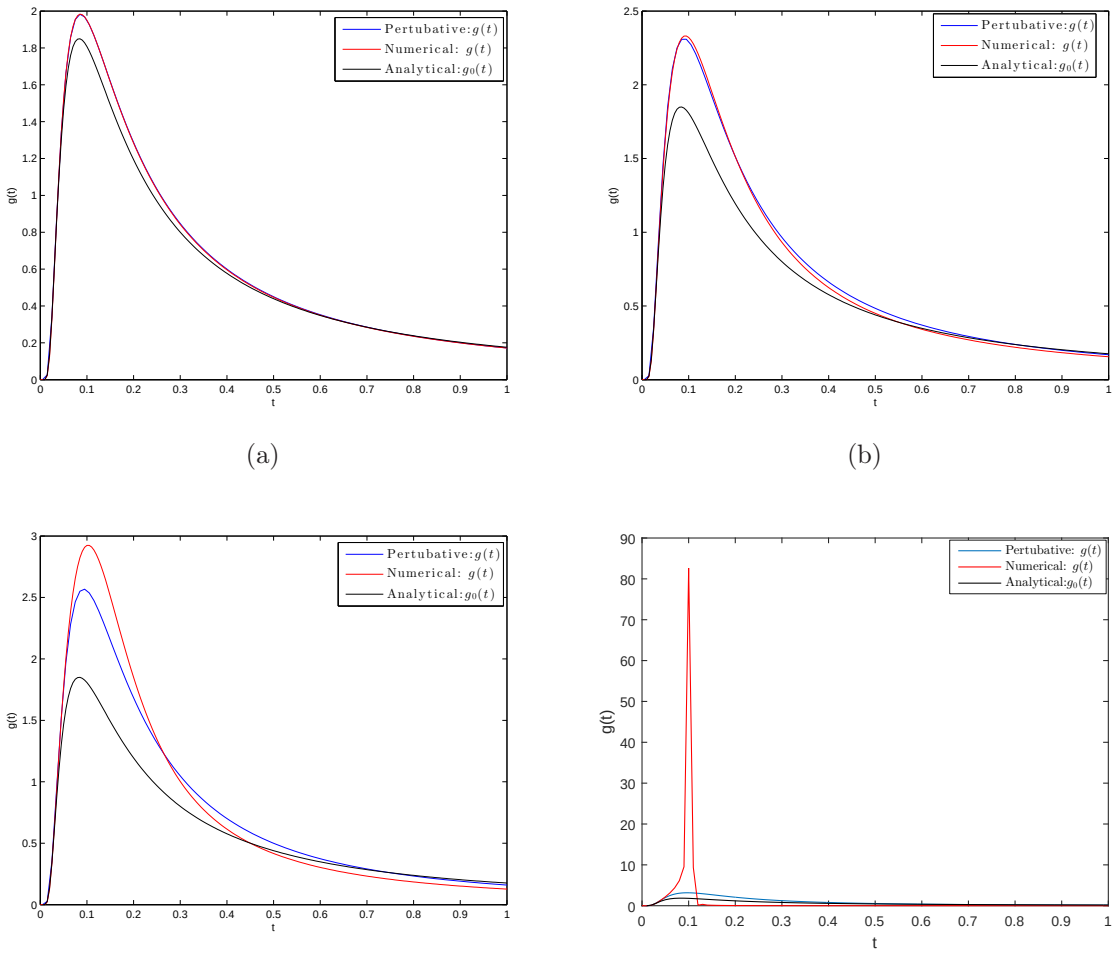
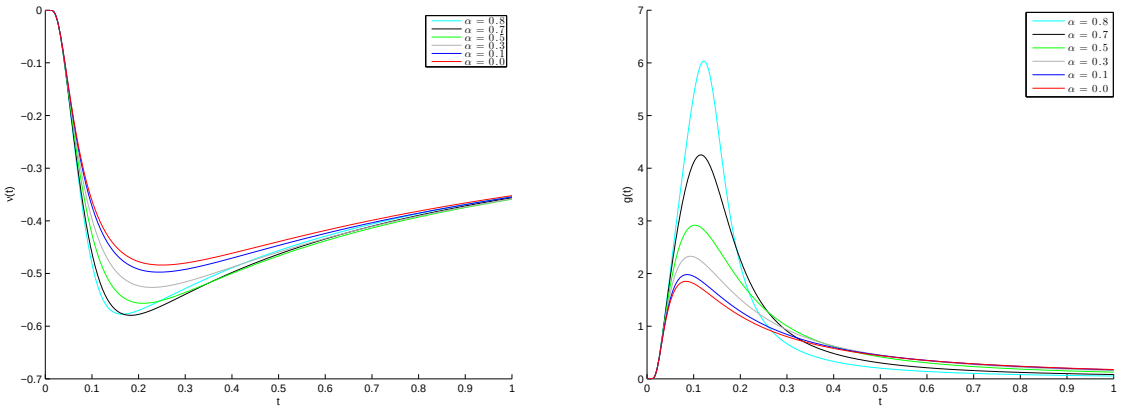
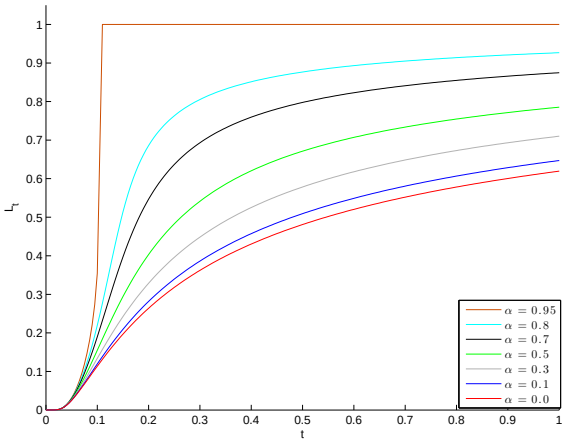


Figure 6.6: Comparison of the numerical and perturbative methods for different values of α.: (a) $\alpha = 0.1$; (b) $\alpha = 0.3$; (c) $\alpha = 0.5$; (d) $\alpha = 1$. For visibility, we take $N = 100$ in the numerical method in the last plot.



(a)

(b)



(c)

Figure 6.7: (a) $\nu(t)$ for different values of α. (b) $g(t)$ for different values of α. (c) $L(t)$ for different values of α computed by solving (6.6) via numerical method.

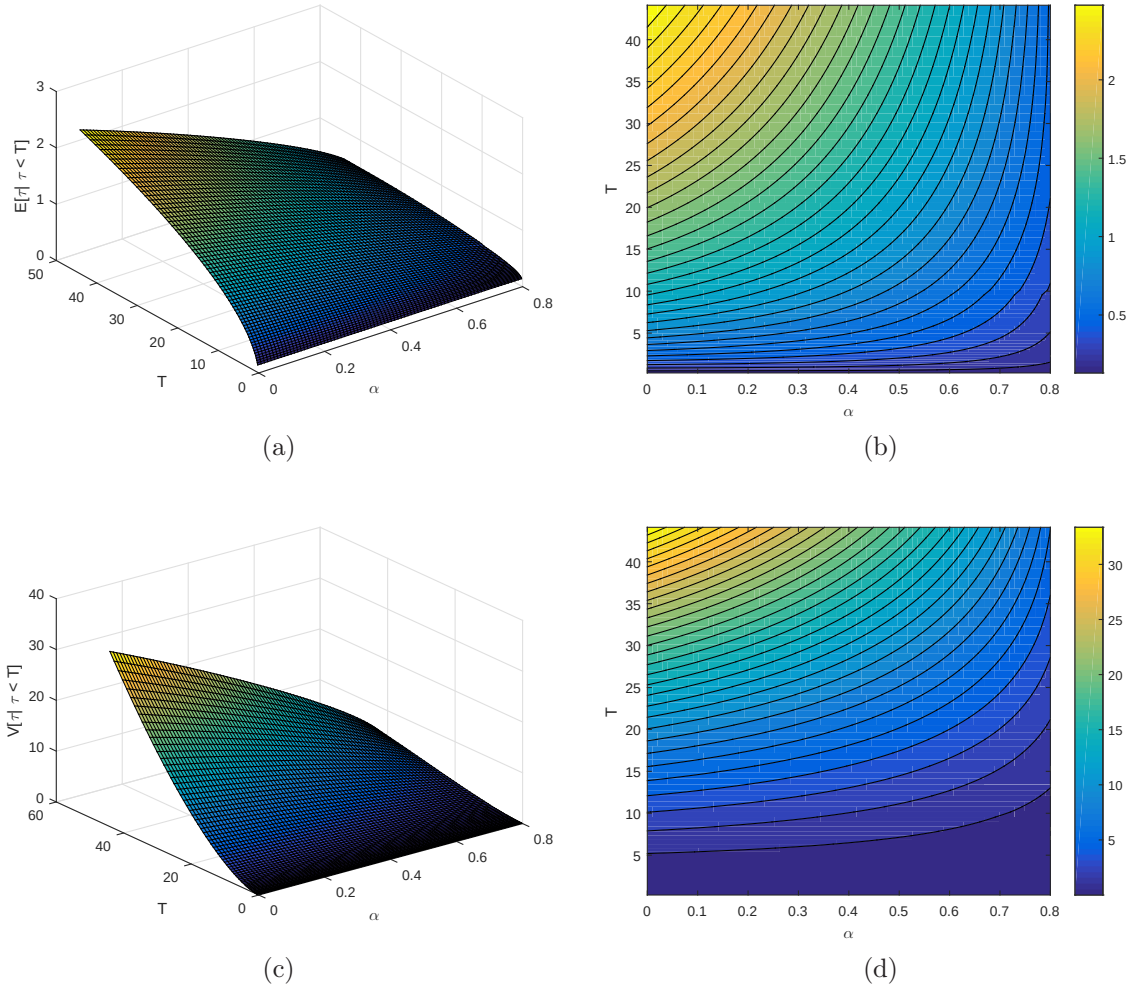


Figure 6.8: (a) and (b): $\mathbb{E}[\tau|\tau < T]$; (c) and (d) $\mathbb{V}[\tau|\tau < T]$ for different values of α and T.

To illustrate the impact of the interaction further, we consider the expectation and variance of the default time depending on α. Since the expectation is infinite, we restrict it to the interval $(0, T)$, i.e. consider $\mathbb{E}[\tau|\tau < T]$ and $\mathbb{V}[\tau|\tau < T]$. These expectations must be finite for any fixed T, and go to infinity when $T \to \infty$. The conditional density is then $p_{\tau|\tau<T}(t) = \frac{p_\tau(t)}{\int_0^T p_\tau(s)\,ds}$ for $t \in [0, T]$. Using this, one can easily evaluate the moments numerically. We present the results in Figure 6.8. As expected, we observe that the expected default time and its variance become smaller with increasing of α, and grow with increasing T.

6.5 Conclusion

We have developed a semi-analytical approach to finding the density of interacting particles where their common downward drift increases in magnitude when particles hit a lower boundary, thus creating a positive feedback effect. This leads to a nonlinear and nonlocal parabolic equation. Using the method of heat potentials, we derived an equivalent

coupled system of Volterra integral equations and solved it numerically, or by expansion for a small interaction parameter α. We confirmed empirically the convergence of order 1 of the numerical method and demonstrated its better complexity in comparison to the particle method in Chapter 5. There are striking financial implications as the computations uncover, in a very simplified setting, how mutual liabilities accelerate defaults of individual banks.

Chapter 7

Conclusions

In this thesis we considered the extended default model with mutual liabilities and measured systemic risk using it. The general case of the model is computationally intense, therefore, we considered several special cases.

For the case when the assets follow a diffusion process, we considered special cases of one, two, and three banks. The equations for one and two banks have analytical solutions which are already known in the literature. We considered the case of three banks, for which we developed a semi-analytical method to compute corresponding transition probabilities. Using this result, we were able to compute the joint survival probability as well as bilateral CVA and DVA semi-analytically.

Next, we developed a numerical method for the case when the assets follow a jump-diffusion process. The method is based on discretization of the integral part of the resulting PIDE and modifying ADI type methods to accommodate it. Unfortunately, it works well only for $N = 2$ and $N = 3$. For a larger N it becomes inefficient and only Monte Carlo methods are considered. Using this numerical method, we computed several important characteristic of the model, such as joint and marginal survival probabilities, CDS and FTD spread, CVA and DVA. We also calibrated model parameters to market quotes and showed the importance of jumps in the model for a short time horizon.

In order to deal with the model for larger N, we considered an infinite-dimensional approximation of the model. We showed that the mean-field limit of the model when the number of banks go to infinity leads to a specific McKean–Vlasov equation, whose drift depends on the absorption rate at the boundary, and analyzed it numerically.

We developed a particle method for the simulation of the process, and proved its convergence of order 1/2 in the time step under certain technical assumptions. Moreover, we improved the convergence rate up to 1 using a Brownian bridge technique and a non-uniform time grid. We confirmed our findings numerically and empirically investigated cases with "blow-up".

We also developed another method for this McKean-Vlasov equation based on the solution of the corresponding PDE for the transition probability. We reduced the problem to a free boundary-value problem, and, using the method of heat potentials, derived a coupled system of Volterra integral equations for the transition density and for the loss through absorption. We developed a numerical method to solve the system as well as an analytical method for the approximation when the value of the interaction parameter is

small. The method has a better convergence rate and allows to compute the loss rate directly.

The mean-field approximation of the model opens up many directions for future research. The nature of the "blow-up" requires more attention, in particular in the proof of the convergence of developed numerical methods. A recent paper by Delarue et al. (2019) reveals new properties of the model, proving existence and uniqueness of the solution after the blow-up. These results can be further used for constructing a global numerical scheme, which guarantees the convergence after the blow up. Another interesting direction is to extend the methods for the model with common noise factor.

Appendix A

Special cases of the default model

In this section, we consider special cases of one, two, and three banks. It gives us a better intuition and understanding of the model and the mechanism of mutual liabilities.

A.1 One bank

In case of just one bank, there are no interbank liabilities, and the model reduces to the standard framework given in Lipton and Sepp (2013). We fully describe it for completeness, since one-dimensional results are used as boundary conditions for higher-dimensional models.

We have assets A and liabilities L, where assets follow jump-diffusion process (2.32)

$$\frac{\mathrm{d}A}{A} = -\kappa\lambda(t)\,\mathrm{d}t + \sigma\,\mathrm{d}W(t) + (\mathrm{e}^J - 1)\,\mathrm{d}N(t), \tag{A.1}$$

where μ is the growth rate, σ the volatility, W is a Brownian motion, N is a Poisson process independent of W, λ is the intensity of jump arrivals, and J is a random jump size with pdf

$$\tilde{\omega}(s) = \begin{cases} 0, & s > 0, \\ \vartheta \mathrm{e}^{\vartheta s}, & s \leq 0, \end{cases} \tag{A.2}$$

with parameters $\vartheta > 0$.

Default boundaries The default occurs at time $\tau = \inf\{t \mid A(t) \leq \Lambda\}$. For the case of one bank, the default boundary (2.41) simplifies to

$$\Lambda = \begin{cases} RL \equiv \Lambda^{<}, & t < T, \\ L \equiv \Lambda^{=}, & t = T, \end{cases} \tag{A.3}$$

where $0 \leq R \leq 1$ is the recovery rate.

Formulation of backward Kolmogorov equation The value function (2.57) for the case of one banks simplifies to

$$V(x,t) = \mathbb{E}^{\mathbb{Q}}\left[\psi(X(T)) \cdot \mathbb{1}_{\{\tau \geq T\}} + \int_t^T \chi(s, X(s)) \cdot \mathbb{1}_{\{\tau > s\}} \mathrm{d}s + \right. \\ \left. + \phi_0(\tau) \cdot \mathbb{1}_{\{\tau < T\}} \middle| X(t) = x\right], \tag{A.4}$$

where $\chi(s,x)$ is the contract payment at an intermediate time $t \le s \le T$ (for example, coupon payment), and $\phi_0(t)$ is the payoffs in case of the intermediate default.

Then, according to the Feynman–Kac formula, the corresponding pricing equation is the Kolmogorov backward equation

$$\frac{\partial V}{\partial t} + \mathcal{L}^{(1)} V = \chi(t,x), \tag{A.5}$$

$$V(t,0) = \phi_0(t), \quad V(t,x) \underset{x\to+\infty}{\longrightarrow} \phi_\infty(t), \tag{A.6}$$

$$V(T,x) = \psi(x), \tag{A.7}$$

with the Kolmogorov backward operator

$$\mathcal{L}^{(1)} f = \frac{1}{2} f_{xx} + \xi f_x + \lambda \mathcal{J} f - \lambda f. \tag{A.8}$$

where

$$\mathcal{J} f(x) = \vartheta \int_0^x f(x) \mathrm{e}^{-\vartheta u} du, \tag{A.9}$$

and ϕ_0, ϕ_∞ are given.

Survival probability The survival probability for one bank is given by

$$Q(t,x) = \mathbb{E}^{\mathbb{Q}}[\mathbb{1}_{\{\tau \ge T, X(T) \ge \mu^=\}} \mid X(t) = x]. \tag{A.10}$$

Then, applying (2.58)–(2.60) with $\psi(x) = \mathbb{1}_{\{x \ge \mu^=\}}$ and $\chi(t,x) = 0$, we get

$$\begin{aligned} &\frac{\partial Q}{\partial t} + \mathcal{L}^{(1)} Q = 0, \\ &Q(t,0) = 0, \\ &Q(T,x) = \mathbb{1}_{\{x \ge \mu^=\}}. \end{aligned} \tag{A.11}$$

Simple CDS In case of one bank, the valuation formula for a CDS simplified to

$$C(t,x) = \mathbb{E}^{\mathbb{Q}} \left[-c \int_t^T \mathbb{1}_{\{\tau > s\}} \, \mathrm{d}s + (1-R) \mathbb{1}_{\{\tau \le T\}} \mid X(t) = x \right]. \tag{A.12}$$

Applying (A.5)–(A.7) with $\chi(t,x) = c$ and $\psi(x) = 1 - R$, we get

$$\begin{aligned} &\frac{\partial C}{\partial t} + \mathcal{L}^{(1)} C = c, \\ &C(t,0) = 1 - R, \\ &C(T,x) = (1-R) \mathbb{1}_{\{x \le \mu^=\}}. \end{aligned} \tag{A.13}$$

A.2 Two banks

In the case of two banks, we have external assets A_1 and A_2, external liabilities L_1 and L_2, as well as interbank mutual liabilities L_{12} and L_{21}, where L_{ij} is the amount the i-th bank owes to the j-th bank. Then, the total assets and liabilities are

$$\begin{aligned} &\tilde{A}_1 = A_1 + L_{21}, \quad \tilde{L}_1 = L_1 + L_{12}, \\ &\tilde{A}_2 = A_2 + L_{12}, \quad \tilde{L}_2 = L_2 + L_{21}. \end{aligned} \tag{A.14}$$

The assets' processes follow (2.32), and the Brownian motions have correlation

$$\mathrm{d}W_1(t)\mathrm{d}W_2(t) = \rho\,\mathrm{d}t. \tag{A.15}$$

We also write the jump processes explicitly. Consider independent Poisson processes $N_{\{1\}}(t), N_{\{2\}}(t)$, and $N_{\{12\}}(t)$ with the corresponding intensities $\lambda_{\{1\}}, \lambda_{\{2\}}$, and $\lambda_{\{12\}}$. Then, we define the processes $N_1(t)$ and $N_2(t)$ as

$$\begin{aligned} N_1(t) &= N_{\{1\}}(t) + N_{\{12\}}(t), \\ N_2(t) &= N_{\{2\}}(t) + N_{\{12\}}(t), \end{aligned} \tag{A.16}$$

with

$$\begin{aligned} \lambda_1 &= \lambda_{\{1\}} + \lambda_{\{12\}}, \\ \lambda_2 &= \lambda_{\{2\}} + \lambda_{\{12\}}. \end{aligned} \tag{A.17}$$

Default boundaries For the case of two banks, the default boundary (2.41) simplifies to

$$\Lambda_i = \begin{cases} R_i(L_i + L_{i\bar{i}}) - L_{\bar{i}i} \equiv \Lambda_i^{<}, & t < T, \\ L_i + L_{i\bar{i}} - L_{\bar{i}i} \equiv \Lambda_i^{=}, & t = T, \end{cases} \tag{A.18}$$

where $0 \le R_i \le 1$ is the recovery rate and $\bar{i} = 3 - i$.

If the k-th bank defaults at intermediate time t', then for the surviving bank $\bar{k} = 3-k$ the default boundary changes to

$$\tilde{\Lambda}_{\bar{k}} = \begin{cases} R_{\bar{k}}(L_{\bar{k}} + L_{\bar{k},k} - R_k L_{k\bar{k}}) \equiv \tilde{\Lambda}_k^{<}, & t < T, \\ L_{\bar{k}} + L_{\bar{k}k} - R_k L_{k\bar{k}} \equiv \tilde{\Lambda}_i^{=}, & t = T. \end{cases} \tag{A.19}$$

It is clear that

$$\Delta\Lambda_k = \tilde{\Lambda}_k - \Lambda_k = \begin{cases} (1 - R_{\bar{k}}R_k)L_{k\bar{k}}, & t < T, \\ (1 - R_k)L_{k\bar{k}}, , & t = T. \end{cases} \tag{A.20}$$

Formulation of backward Kolmogorov equation The value function (2.57) for the case of two banks simplifies to

$$\begin{aligned} V(x,t) = \mathbb{E}^{\mathbb{Q}} \Big[\psi(X(T)) \cdot \mathbb{1}_{\{\tau \ge T\}} + \int_t^T \chi(s, X(s)) \cdot \mathbb{1}_{\{\tau > s\}} \mathrm{d}s + \\ + \phi_{1,0}(\tau_1, X_2(\tau_1)) \cdot \mathbb{1}_{\{\tau_1 < T, \tau_1 = \tau\}} + \phi_{2,0}(\tau_2, X_1(\tau_2)) \cdot \mathbb{1}_{\{\tau_2 < T, \tau_2 = \tau\}} \big| \, X(t) = x \Big]. \end{aligned} \tag{A.21}$$

Accordingly, (2.58)–(2.60) becomes

$$\frac{\partial V}{\partial t} + \mathcal{L}^{(2)} V = \chi(t,x), \tag{A.22}$$

$$V(t,0,x_2) = \phi_{1,0}(t,x_2), \quad V(t,x_1,x_2) \underset{x_1 \to +\infty}{\longrightarrow} \phi_{1,\infty}(t,x_2), \tag{A.23}$$

$$V(t,x_1,0) = \phi_{2,0}(t,x_1), \quad V(t,x_1,x_2) \underset{x_2 \to +\infty}{\longrightarrow} \phi_{2,\infty}(t,x_1), \tag{A.24}$$

$$V(T,x) = \psi(x), \tag{A.25}$$

where Kolmogorov backward operator is

$$\mathcal{L}^{(2)}f = \frac{1}{2}f_{x_1x_1} + \rho f_{x_1x_2} + \frac{1}{2}f_{x_2x_2} + \xi_1 f_{x_1} + \xi_2 f_{x_2} + \lambda_1\mathcal{J}_1 f + \lambda_2\mathcal{J}_2 f + \lambda_{12}\mathcal{J}_{12}f - vf = \\ = \Delta_\rho f + \xi\cdot\nabla f + \mathcal{J}f - vf, \quad \text{(A.26)}$$

where $v = \lambda_1 + \lambda_2 + \lambda_{12}$ and

$$\mathcal{J}_1 f(x) = \varsigma_1 \int_0^{x_1} f(x_1 - u, x_2)\mathrm{e}^{-\varsigma_1 u}\mathrm{d}u, \quad \text{(A.27)}$$

$$\mathcal{J}_2 f(x) = \varsigma_2 \int_0^{x_2} f(x_1, x_2 - u)\mathrm{e}^{-\varsigma_2 u}d\mathrm{d}u, \quad \text{(A.28)}$$

$$\mathcal{J}_{12} f(x) = \mathcal{J}_1\mathcal{J}_2 f(x) = \varsigma_1\varsigma_2 \int_0^{x_1}\int_0^{x_2} f(x_1 - u, x_2 - v)\mathrm{e}^{-\varsigma_1 u - \varsigma_2 v}\mathrm{d}u\mathrm{d}v. \quad \text{(A.29)}$$

Joint survival probability The joint survival probability simplifies to

$$Q(t,x) = \mathbb{E}^{\mathbb{Q}}[\mathbb{1}_{\{\tau\geq T, X_1(T)\geq\bar{\mu}_1^{=}, X_2(T)\geq\bar{\mu}_2^{=}\}} \mid X(t) = x]. \quad \text{(A.30)}$$

Accordingly,

$$\begin{aligned} &\frac{\partial Q}{\partial t} + \mathcal{L}Q = 0, \\ &Q(t, x_1, 0) = 0, \quad Q(t, 0, x_2) = 0, \\ &Q(T, x_1, x_2) = \mathbb{1}_{\{x_1\geq\bar{\mu}_1^{=}, x_2\geq\bar{\mu}_2^{=}\}}. \end{aligned} \quad \text{(A.31)}$$

Marginal survival probability The marginal survival probability for the first bank simplifies to

$$q_1(t,x) = \mathbb{E}^{\mathbb{Q}}[\mathbb{1}_{\{\tau\geq T, X(T)\in D_1\cup D_{12})\}} + \Xi(\tau_2, X_1(\tau_2))\cdot\mathbb{1}_{\{\tau_2<T, \tau_2=\tau\}} \mid X(t) = x], \quad \text{(A.32)}$$

where D_{12} is the set where both banks survive at the terminal time, D_1 is the set where only the first bank survives, and $\Xi(\tau_2, X_1(\tau_2))$ is the one-dimensional survival probability with the modified boundaries.

Then, applying (2.58)–(2.60) with $\psi(x) = \mathbb{1}_{\{x\in D_1\cup D_{12})\}}$, $\chi(t,x) = 0$, we get

$$\begin{aligned} &\frac{\partial}{\partial t}q_1(t,x) + \mathcal{L}q_1(t,x) = 0, \\ &q_1(t,0,x_2) = 0, \quad q_1(t,x_1,0) = \Xi(t,x_1) = \begin{cases}\chi_{1,0}(t,x_1), & x_1\geq\tilde{\mu}_1^{<}, \\ 0, & x_1<\tilde{\mu}_1^{<},\end{cases} \\ &q_1(t,x_1,x_2) \underset{x_1\to+\infty}{\longrightarrow} 1, \quad q_1(t,x_1,x_2) \underset{x_2\to+\infty}{\longrightarrow} \chi_{1,\infty}(t,x_1), \\ &q_1(T,x) = \mathbb{1}_{\{x\in D_1\cup D_{12}\}}, \end{aligned} \quad \text{(A.33)}$$

where $\chi_{1,0}(t,x_1)$ and $\chi_{1,\infty}(t,x_1)$ are the survival probabilities for one bank defined in Section A.1.

Credit default swap In the case of two banks, the CDS pricing equation (2.70) simplifies to

$$C_1(t,x) = \mathbb{E}^{\mathbb{Q}}\Big[-c\int_t^T \mathbb{1}_{\{\tau>s\}}\,\mathrm{d}s + (1-\min(R_1,\tilde{R}_1)\mathbb{1}_{\{X(T)\in\hat{D}\cup D_2\}} \\ +(1-R_1)\mathbb{1}_{\{\tau_1<T,\tau_1=\tau\}} + c_{1,0}(\tau_2,X_1(\tau_2))\mathbb{1}_{\{\tau_2<T,\tau_2=\tau\}} \mid X(t)=x\Big], \quad \text{(A.34)}$$

and (2.73) transforms to

$$\begin{aligned}
&\frac{\partial}{\partial t}C_1(t,x) + \mathcal{L}^{(2)}C_1(t,x) = c,\\
&C_1(t,0,x_2) = 1-R_1, \quad C_1(t,x_1,x_2) \underset{x_1\to+\infty}{\longrightarrow} -c(T-t),\\
&C_1(t,x_1,0) = \begin{cases} c_{1,0}(t,x_1), & x_1\geq\tilde{\mu}_1,\\ 1-R_1, & x_1<\tilde{\mu}_i,\end{cases} \quad C_1(t,x_1,x_2)) \underset{x_2\to+\infty}{\longrightarrow} c_{1,\infty}(t,x_1),\\
&C_1(T,x) = \psi(x) = \begin{cases} 1-\min(R_1,\tilde{R}_1(1)), & (x_1,x_2)\in D_2,\\ 1-\min(R_1,\tilde{R}_1(\omega_2)), & (x_1,x_2)\in D_{12},\end{cases}
\end{aligned} \quad \text{(A.35)}$$

where

$$\tilde{R}_1(\omega_2) = \min\left[1, \frac{A_1(T)+\omega_2 L_{21}(T)}{L_1(T)+\omega_2 L_{12}(T)}\right],$$

where $c_{1,0}(t,x_1)$ and $c_{1,\infty}(t,x_1)$ are the solutions of one-dimensional problems described in Section A.1.

First-to-default swap In the case of two banks, the FTD pricing equation (2.74) simplifies to

$$F(t,x) = \mathbb{E}^{\mathbb{Q}}\Big[-c\int_t^T \mathbb{1}_{\{\tau>s\}}\,\mathrm{d}s + \beta_0\mathbb{1}_{\{X(T)\in\hat{D}\}} + \beta_1\mathbb{1}_{\{X(T)\in D_1\}} + \beta_2\mathbb{1}_{\{X(T)\in D_2\}} \\ +(1-R_1)\mathbb{1}_{\{\tau_1<T,\tau_1=\tau\}} + (1-R_2)\mathbb{1}_{\{\tau_2<T,\tau_2=\tau\}} \mid X(t)=x\Big], \quad \text{(A.36)}$$

where

$$\beta_1 = 1-\min(R_2,\tilde{R}_2(\omega)), \quad \beta_2 = 1-\min(R_1,\tilde{R}_1(\omega)),\\ \beta_0 = 1-\min(\min(\tilde{R}_1(\omega),R_1),\min(\tilde{R}_2(\omega),R_2)),$$

and (2.75) transforms to

$$\begin{aligned}
&\frac{\partial}{\partial t}F(t,x) + \mathcal{L}^{(2)}F(t,x) = c,\\
&F(t,0,x_2) = 1-R_1, \quad F(t,x_1,0) = 1-R_2,\\
&F(t,x_1,x_2) \underset{x_2\to+\infty}{\longrightarrow} f_{1,\infty}(t,x_1), \quad F(t,x_1,x_2) \underset{x_1\to+\infty}{\longrightarrow} f_{2,\infty}(t,x_2),\\
&F(T,x) = \beta_0\mathbb{1}_{\{x\in\hat{D}\}} + \beta_1\mathbb{1}_{\{x\in D_1\}} + \beta_2\mathbb{1}_{\{x\in D_2\}},
\end{aligned} \quad \text{(A.37)}$$

where $f_{1,\infty}(t,x_1)$ and $f_{2,\infty}(t,x_2)$ are the solution of one-dimensional CDS pricing problems described in Section A.2.

Credit and Debt Value Adjustments for CDS When we have two banks, we consider unilateral CVA and DVA described in Section 2.1.5.

For CVA, we associate x_1 with the Protection Seller and x_2 with the Reference Name. Then, using (2.24), it can be given by the solution of (2.58)–(2.60) with $\chi(t,x) = 0$ and $\psi(x) = 0$,

$$\begin{aligned} &\frac{\partial}{\partial t}V^{CVA} + \mathcal{L}^{(2)}V^{CVA} = 0, \\ &V^{CVA}(t,0,x_2) = (1-R_{PS})V^{CDS}(t,x_2)^+, \quad V^{CVA}(t,x_1,0) = 0, \\ &V^{CVA}(T,x_1,x_2) = 0. \end{aligned} \tag{A.38}$$

Similarly, for DVA, we associate x_1 with the Protection Buyer and x_2 with the Reference Name. Then, using (2.26), DVA can be given by the solution of (2.58)–(2.60),

$$\begin{aligned} &\frac{\partial}{\partial t}V^{DVA} + \mathcal{L}^{(2)}V^{DVA} = 0, \\ &V^{DVA}(t,0,x_2) = (1-R_{PB})V^{CDS}(t,x_2)^-, \quad V^{DVA}(t,x_1,0) = 0, \\ &V^{DVA}(T,x_1,x_2) = 0. \end{aligned} \tag{A.39}$$

A.3 Three banks

In contrast to the previous section, it is not so informative to write default boundaries explicitly. Thus, we just use Sections 2.2 for $N = 3$. We concentrate on the formulation of Kolmogorov backward equations for joint survival probability and CVA/DVA, which we use in the following chapters.

For the joint survival probability, (2.58)–(2.60) transforms to

$$\frac{\partial Q}{\partial t} + \mathcal{L}^{(3)}Q = 0, \tag{A.40}$$

$$Q(t,0,x_2,x_3) = 0, \quad Q(t,x_1,0,x_3) = 0, \quad Q(t,x_1,x_2,0) = 0, \tag{A.41}$$

$$Q(T,x_1,x_2,x_3) = \mathbb{1}_{\{x_1 \geq \mu_1^=, x_2 \geq \mu_2^=, x_3 \geq \mu_3^=\}}. \tag{A.42}$$

For CVA, we associate x_1 with the Protection Seller (PS), x_2 with the Protection Buyer (PB), and x_3 with the Reference Name (RN). Then, (2.77) can be rewritten as

$$\begin{aligned} &\frac{\partial}{\partial t}V^{CVA} + \mathcal{L}^{(3)}V^{CVA} = 0, \\ &V^{CVA}(t,0,x_2,x_3) = (1-R_1)V^{CDS}(t,x_3)_+, \\ &V^{CVA}(t,x_1,0,x_3) = 0, \quad V^{CVA}(t,x_1,x_2,0) = 0, \\ &V^{CVA}(T,x_1,x_2,x_3) = 0. \end{aligned} \tag{A.43}$$

Similarly, for DVA, (2.79) can be rewritten as

$$\begin{aligned} &\frac{\partial}{\partial t}V^{DVA} + \mathcal{L}^{(3)}V^{DVA} = 0, \\ &V^{DVA}(t,0,x_2,x_3) = (1-R_2)V^{CDS}(t,x_3)_-, \\ &V^{DVA}(t,x_1,0,x_3) = 0, \quad V^{DVA}(t,x_1,x_2,0) = 0, \\ &V^{DVA}(T,x_1,x_2,x_3) = 0. \end{aligned} \tag{A.44}$$

Appendix B

The method of heat potentials

We follow Tikhonov and Samarskii (1963), Section 6.4 in this section.

Consider a heat equation

$$u_t = a^2 u_{xx}, \tag{B.1}$$

with moving boundaries $\chi_1(t)$ and $\chi_2(t)$ and zero initial condition. The boundary conditions are

$$u(t, \chi_1(t)) = g_1(t), \tag{B.2}$$
$$u(t, \chi_2(t)) = g_2(t). \tag{B.3}$$

First, we consider a solution of (B.1) without the boundary conditions. It is well-known that the Green's function for (B.1) on $\mathbb{R}$ is

$$G(\tau, x', x) = \frac{1}{2\sqrt{\pi a^2 \tau}} e^{\frac{(x'-x)^2}{4a^2\tau}}. \tag{B.4}$$

B.1 Heat potentials

Consider two integrals

$$V = a^2 \int_0^t G|_{x'=\chi_1(\tau)}\, \nu \,\mathrm{d}\tau, \tag{B.5}$$
$$W = 2a^2 \int_0^t \left.\frac{\partial G}{\partial x'}\right|_{x'=\chi_1(\tau)} \mu \,\mathrm{d}\tau, \tag{B.6}$$

for some $\nu(t)$ and $\mu(t)$.

Using (B.4), the last two equations become

$$V = \frac{a^2}{2\sqrt{\pi}} \int_0^t \frac{1}{\sqrt{a^2(t-\tau)}} \mathrm{e}^{-\frac{(x-\chi_1(\tau))^2}{4a^2(t-\tau)}} \nu(\tau)\,\mathrm{d}\tau, \tag{B.7}$$
$$W = \frac{1}{2a\sqrt{\pi}} \int_0^t \frac{x-\chi_1(\tau)}{(t-\tau)^{3/2}} \mathrm{e}^{-\frac{(x-\chi_1(\tau))^2}{4a^2(t-\tau)}} \mu(\tau)\,\mathrm{d}\tau. \tag{B.8}$$

Functions V and W are called single layer and double layer heat potentials. In the rest of the section, we study the behaviour of the heat potentials along the curve $x = \chi(t)$.

We want to show that

$$W|_{x=\chi_1(t)+0} = W|_{x=\chi_1(t)} + \mu(t), \tag{B.9}$$
$$W|_{x=\chi_1(t)-0} = W|_{x=\chi_1(t)} - \mu(t), \tag{B.10}$$

and

$$\left.\frac{\partial V}{\partial x}\right|_{x=\chi_1(t)+0} = \left.\frac{\partial V}{\partial x}\right|_{x=\chi_1(t)} + \frac{\nu(t)}{2}, \tag{B.11}$$
$$\left.\frac{\partial V}{\partial x}\right|_{x=\chi_1(t)-0} = \left.\frac{\partial V}{\partial x}\right|_{x=\chi_1(t)} - \frac{\nu(t)}{2}, \tag{B.12}$$

assuming that $\chi_1(t)$, $\mu(t)$, and $\nu(t)$ are differentiable.

Consider W^0, defined with constant $\mu(t) = \mu_0$:

$$W^0 = \frac{1}{2a\sqrt{\pi}} \int_0^t \frac{x - \chi_1(\tau)}{(t-\tau)^{3/2}} \mathrm{e}^{-\frac{(x-\chi_1(\tau))^2}{4a^2(t-\tau)}} \mu_0 \,\mathrm{d}\tau, \tag{B.13}$$

and auxiliary function

$$\tilde{V}^0 = \frac{1}{2a\sqrt{\pi}} \int_0^t \frac{2\chi_1'(\tau)}{\sqrt{t-\tau}} \mathrm{e}^{-\frac{(x-\chi_1(\tau))^2}{4a^2(t-\tau)}} \mu_0 \,\mathrm{d}\tau. \tag{B.14}$$

The difference $W^0 - \tilde{V}^0$ is

$$\begin{aligned} W^0(t,x) - \tilde{V}^0(t,x) &= \frac{1}{2a\sqrt{\pi}} \int_0^t \mathrm{e}^{-\frac{(x-\chi_1(\tau))^2}{4a^2(t-\tau)}} \left[\frac{x-\chi_1(\tau)}{(t-\tau)^{3/2}} - \frac{2\chi_1'(\tau)}{\sqrt{t-\tau}}\right] \mu_0 \,\mathrm{d}\tau \\ &= \frac{2\mu_0}{\sqrt{\pi}} \int_{\frac{x-\chi_1(0)}{2a\sqrt{t}}}^{\alpha_0} \mathrm{e}^{-\alpha^2} \,\mathrm{d}\alpha, \end{aligned}$$

where

$$\alpha = \frac{x - \chi_1(\tau)}{2a\sqrt{t-\tau}},$$
$$\alpha_0 = \begin{cases} +\infty, & x > \chi_1(t), \\ 0, & x = \chi_1(t), \\ -\infty, & x < \chi_1(t). \end{cases}$$

Hence,

$$[W^0(t,x_0 \pm 0) - W^0(t,x_0)] - [\tilde{V}^0(t,x_0 \pm 0) - \tilde{V}^0(t,x_0)] = \frac{2\mu_0}{\sqrt{\pi}} \int_0^{\pm\infty} \mathrm{e}^{-\alpha^2} \mathrm{d}\alpha = \pm\mu_0. \tag{B.15}$$

Since $\tilde{V}$ is continuous, we have

$$W^0(t, x_0 \pm 0) = W^0(t, x_0) \pm \mu_0. \tag{B.16}$$

As a result

$$W(t,x) = W^0(t,x) - \psi(t,x), \tag{B.17}$$

where

$$\psi(t,x) = \frac{a^2}{2\sqrt{\pi}} \int_0^t \frac{x-\chi_1(\tau)}{[a^2(t-\tau)]^{3/2}} \mathrm{e}^{-\frac{(x-\chi_1(\tau))^2}{4a^2(t-\tau)}} \left[\mu(t)-\mu(\tau)\right] \mathrm{d}\tau.$$

Since we assumed that $\mu(t)$ is differentiable, the integral in the last equation has a square-root (integrable) singularity, and, therefore $\psi(t,x)$ is continuous. Thus,

$$W(t, x_0 \pm 0) = W(t, x_0) \pm \mu(t), \tag{B.18}$$

and we get (B.9)–(B.10).

Similarly, taking the derivative $\frac{\partial V}{\partial x}$,

$$\frac{\partial V}{\partial x} = -\frac{1}{2a\sqrt{\pi}} \frac{1}{2} \int_0^t \frac{x-\chi_1(\tau)}{(t-\tau)^{3/2}} \mathrm{e}^{-\frac{(x-\chi_1(\tau))^2}{4a^2(t-\tau)}} \mu(\tau)\, \mathrm{d}\tau, \tag{B.19}$$

and is equal to $-W(t,x)$ with density $\mu(t) = \frac{\nu(t)}{2}$.

As a result, we immediately get (B.11)–(B.12).

B.2 Boundary value problem

In this section we show how to solve (B.1) with boundary conditions (B.2) (assuming $\chi_2(t) = +\infty$ in (B.3)), and the initial condition $u(0,x) = 0$.

Using Green's theorem, we can write the solution as

$$u(t,x) = \frac{1}{2a^2} W(t,x) = \int_0^t \frac{\partial G}{\partial x'}(t-\tau, x, \chi_1(\tau)\mu(\tau))\, \mathrm{d}\tau. \tag{B.20}$$

Using the expression for the Green's function, we have

$$u(t,x) = \frac{1}{4\sqrt{\pi}} \int_0^t \frac{x-\chi_1(t)}{[a^2(t-\tau]^{3/2}} \mathrm{e}^{-\frac{(x-\chi_1(\tau))^2}{4a^2(t-\tau)}} \mu(\tau)\, \mathrm{d}\tau. \tag{B.21}$$

This function satisfies (B.1) for $x > \chi_1(t)$, has zero initial condition, and bounded for any choice of $\mu(t)$. Moreover, at the point $x = \chi_1(t)$, as we showed in the previous section, the function is discontinuous, and using (B.18), we have

$$\frac{\mu(t)}{2a^2} + \frac{1}{4\sqrt{\pi}} \int_0^t \frac{\chi_1(t)-\chi_1(\tau}{[a^2(t-\tau)]^{3/2}} \mathrm{e}^{-\frac{(\chi_1(t)-\chi_1(\tau))^2}{4a^2(t-\tau)}} \mu(\tau)\, \mathrm{d}\tau = g_1(t). \tag{B.22}$$

The last equation is a Volterra integral equation of the second kind. Solving this equation, we find the density function $\mu(t)$, and therefore can find the solution of the original problem in the form (B.21).

Bibliography

Abate, J., Choudhury, G. L., and Whitt, W. (2000). An introduction to numerical transform inversion and its application to probability models. In *Computational Probability*, pages 257–323. Springer.

Abramowitz, M. and Stegun, I. A. (1964). *Handbook of Mathematical Functions: with formulas, graphs, and mathematical tables*, volume 55. Courier Corporation.

Albanese, C. and Chen, O. X. (2005). Discrete credit barrier models. *Quantitative Finance*, 5(3):247–256.

Allen, F. and Gale, D. (2000). Financial contagion. *Journal of Political Economy*, 108(1):1–33.

Andersen, L. and Andreasen, J. (2000). Jump-diffusion processes: Volatility smile fitting and numerical methods for option pricing. *Review of Derivatives Research*, 4(3):231–262.

Antonelli, F., Kohatsu-Higa, A., et al. (2002). Rate of convergence of a particle method to the solution of the McKean–Vlasov equation. *The Annals of Applied Probability*, 12(2):423–476.

Artzner, P. and Delbaen, F. (1995). Default risk insurance and incomplete markets. *Mathematical Finance*, 5(3):187–195.

Asmussen, S., Glynn, P., and Pitman, J. (1995). Discretization error in simulation of one-dimensional reflecting Brownian Motion. *The Annals of Applied Probability*, 5(4):875–896.

Avellaneda, M. and Zhu, J. (2001). Distance to default. *Risk*, 14(12):125–129.

Ballotta, L., Fusai, G., and Marazzina, D. (2019). Integrated structural approach to credit value adjustment. *European Journal of Operational Research*, 272(3):1143–1157.

Barker, R., Dickinson, A., Lipton, A., and Virmani, R. (2016). Systemic risks in CCP networks. *arXiv preprint arXiv:1604.00254*.

Beyn, W.-J. (2012). An integral method for solving nonlinear eigenvalue problems. *Linear Algebra and its Applications*, 436(10):3839–3863.

Bielecki, T. R. and Rutkowski, M. (2013). *Credit risk: modeling, valuation and hedging*. Springer Science & Business Media.

Black, F. and Cox, J. C. (1976). Valuing corporate securities: Some effects of bond indenture provisions. *The Journal of Finance*, 31(2):351–367.

Boen, L. and Hout, K. J. (2019). Operator splitting schemes for the two-asset merton jump-diffusion model. *arXiv preprint arXiv:1901.03839*.

Borovkov, K. and Novikov, A. (2005). Explicit bounds for approximation rates of boundary crossing probabilities for the Wiener process. *Journal of Applied Probability*, 42(1):82–92.

Bossy, M. and Talay, D. (1997). A stochastic particle method for the McKean-Vlasov and the Burgers equation. *Mathematics of Computation*, 66(217):157–192.

Brunner, H. (1985). The numerical solution of weakly singular Volterra integral equations by collocation on graded meshes. *Mathematics of Computation*, 45(172):417–437.

Bush, N., Hambly, B. M., Haworth, H., Jin, L., and Reisinger, C. (2011). Stochastic evolution equations in portfolio credit modelling. *SIAM Journal on Financial Mathematics*, 2(1):627–664.

Butcher, J. C. (2008). *Numerical Methods for Ordinary Differential Equations*. John Wiley & Sons.

Cáceres, M. J., Carrillo, J. A., and Perthame, B. (2011). Analysis of nonlinear noisy integrate & fire neuron models: blow-up and steady states. *The Journal of Mathematical Neuroscience*, 1(1):7.

Cariboni, J. and Schoutens, W. (2007). Pricing credit default swaps under Lévy models. *Journal of Computational Finance*, 10(4):71.

Carmona, R. and Delarue, F. (2018). *Probabilistic Theory of Mean Field Games with Applications I-II*. Springer.

Carrillo, J. A., González, M. d. M., Gualdani, M. P., and Schonbek, M. E. (2013). Classical solutions for a nonlinear Fokker–Planck equation arising in computational neuroscience. *Communications in Partial Differential Equations*, 38(3):385–409.

Charney, J. G., Fjörtoft, R., and Neumann, J. v. (1950). Numerical integration of the barotropic vorticity equation. *Tellus*, 2(4):237–254.

Conn, A. R., Gould, N. I., and Toint, P. L. (2000). *Trust region methods*, volume 1. Siam.

Cont, R. and De Larrard, A. (2012). Order book dynamics in liquid markets: limit theorems and diffusion approximations. *Working paper*.

Cont, R., Moussa, A., and Santos, E. (2010). Network structure and systemic risk in banking systems. *Available at SSRN: https://ssrn.com/abstract=1733528 or http://dx.doi.org/10.2139/ssrn.1733528*.

David, A. and Lehar, A. (2014). Why are banks highly interconnected? *Available at SSRN 1108870*.

Delarue, F., Inglis, J., Rubenthaler, S., and Tanré, E. (2015a). Global solvability of a networked integrate-and-fire model of McKean–Vlasov type. *The Annals of Applied Probability*, 25(4):2096–2133.

Delarue, F., Inglis, J., Rubenthaler, S., and Tanré, E. (2015b). Particle systems with a singular mean-field self-excitation. Application to neuronal networks. *Stochastic Processes and their Applications*, 125(6):2451–2492.

Delarue, F., Nadtochiy, S., and Shkolnikov, M. (2019). Global solutions to the supercooled Stefan problem with blow-ups: regularity and uniqueness. *arXiv preprint arXiv:1902.05174*.

Demmel, J., Dumitriu, I., and Holtz, O. (2007). Fast linear algebra is stable. *Numerische Mathematik*, 108(1):59–91.

Driscoll, T. A., Hale, N., and Trefethen, L. N. (2014). Chebfun guide. *Pafnuty Publ.*

Duffie, D. and Singleton, K. J. (1997). An econometric model of the term structure of interest-rate swap yields. *The Journal of Finance*, 52(4):1287–1321.

Duffie, D. and Singleton, K. J. (1999). Modeling term structures of defaultable bonds. *Review of Financial studies*, 12(4):687–720.

Eisenberg, L. and Noe, T. H. (2001). Systemic risk in financial systems. *Management Science*, 47(2):236–249.

Escobar, M., Ferrando, S., and Wen, X. (2013). Three dimensional distribution of Brownian motion extrema. *Stochastics: An International Journal of Probability and Stochastic Processes*, 85(5):807–832.

Escobar, M., Ferrando, S., and Wen, X. (2014). Barrier options in three dimensions. *International Journal of Financial Markets and Derivatives*, 3(3):260–292.

Feng, L. and Linetsky, V. (2008). Pricing discretely monitored barrier options and defaultable bonds in Lévy process models: a fast Hilbert transform approach. *Mathematical Finance*, 18(3):337–384.

Finger, C., Finkelstein, V., Pan, G., Lardy, J.-P., Ta, T., and Tierney, J. (2002). Creditgrades technical document. *RiskMetrics Group*, pages 1–51.

Flügge, S. (2012). *Practical quantum mechanics*. Springer Science & Business Media.

Gai, P., Haldane, A., and Kapadia, S. (2011). Complexity, concentration and contagion. *Journal of Monetary Economics*, 58(5):453–470.

Gai, P. and Kapadia, S. (2010). Contagion in financial networks. In *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, page rspa20090410. The Royal Society.

Giesecke, K., Spiliopoulos, K., Sowers, R. B., et al. (2013). Default clustering in large portfolios: Typical events. *The Annals of Applied Probability*, 23(1):348–385.

Glasserman, P. (2013). *Monte Carlo methods in financial engineering*, volume 53. Springer Science & Business Media.

Glasserman, P. and Young, H. P. (2016). Contagion in financial networks. *Journal of Economic Literature*, 54(3):779–831.

Gobet, E. and Pagliarani, S. (2018). Analytical approximations of non-linear SDEs of McKean–Vlasov type. *Journal of Mathematical Analysis and Applications*, 466(1):71–106.

Gradštejn, I. S., Ryžik, I. M., Jeffrey, A., and Zwillinger, D. (2007). *Table of integrals, series, and products*. Elsevier/Academic Press.

Guibas, L. and Stolfi, J. (1985). Primitives for the manipulation of general subdivisions and the computation of voronoi. *ACM Transactions on Graphics (TOG)*, 4(2):74–123.

Hairer, E., Lubich, C., and Schlichte, M. (1985). Fast numerical solution of nonlinear Volterra convolution equations. *SIAM Journal on Scientific and Statistical Computing*, 6(3):532–541.

Hambly, B., Ledger, S., and Sojmark, A. (2018). A McKean–Vlasov equation with positive feedback and blow-ups. *arXiv preprint arXiv:1801.07703*.

Hambly, B. and Sojmark, A. (2018). An SPDE model for systemic risk with endogenous contagion. *arXiv preprint arXiv:1801.10088*.

Haworth, H. (2006). *Structural models of credit with default contagion*. PhD thesis, University of Oxford.

Haworth, H., Reisinger, C., and Shaw, W. (2008). Modelling bonds and credit default swaps using a structural model with contagion. *Quantitative Finance*, 8(7):669–680.

He, H., Keirstead, W. P., and Rebholz, J. (1998). Double lookbacks. *Mathematical Finance*, 8(3):201–228.

Hilberink, B. and Rogers, L. (2002). Optimal capital structure and endogenous default. *Finance and Stochastics*, 6(2):237–263.

Hull, J. C. and White, A. (2001). Valuing credit default swaps II: modeling default correlations. *Journal of Derivatives*, 8(3):12–22.

Hundsdorfer, W. and Verwer, J. G. (2013). *Numerical solution of time-dependent advection-diffusion-reaction equations*, volume 33. Springer Science & Business Media.

Hurd, T. R. (2016). *Contagion! Systemic Risk in Financial Networks*. Springer.

Hyer, T., Lipton, A., Pugachevsky, D., and Qui, S. (1999). A hidden-variable model for risky bonds. Technical report, working paper, Bankers Trust.

Ichiba, T., Ludkovski, M., and Sarantsev, A. (2018). Dynamic contagion in a banking system with births and defaults. *arXiv preprint arXiv:1807.09897*.

In't Hout, K. and Foulon, S. (2010). ADI finite difference schemes for option pricing in the Heston model with correlation. *International Journal of Numerical Analysis and Modeling*, 7(2):303–320.

In't Hout, K. and Welfert, B. (2007). Stability of ADI schemes applied to convection–diffusion equations with mixed derivative terms. *Applied Numerical Mathematics*, 57(1):19–35.

in't Hout, K. J. and Toivanen, J. (2018). Adi schemes for valuing european options under the bates model. *Applied Numerical Mathematics*, 130:143–156.

Itkin, A. and Carr, P. (2011). Jumps without tears: A new splitting technology for barrier options. *International Journal of Numerical Analysis and Modeling*, 8(4):667–704.

Itkin, A. and Lipton, A. (2015). Efficient solution of structural default models with correlated jumps and mutual obligations. *International Journal of Computer Mathematics*, 92(12):2380–2405.

Itkin, A. and Lipton, A. (2017). Structural default model with mutual obligations. *Review of Derivatives Research*, 20:15–46.

Jarrow, R. A. and Turnbull, S. M. (1995). Pricing derivatives on financial securities subject to credit risk. *The Journal of Finance*, 50(1):53–85.

Javed, M. and Trefethen, L. N. (2014). A trapezoidal rule error bound unifying the Euler–Maclaurin formula and geometric convergence for periodic functions. In *Proceedings of the Royal Society A*, volume 470:2161.

Kaushansky, V., Lipton, A., and Reisinger, C. (2018a). Numerical analysis of an extended structural default model with mutual liabilities and jump risk. *Journal of Computational Science*, 24:218–231.

Kaushansky, V., Lipton, A., and Reisinger, C. (2018b). Transition probability of Brownian motion in the octant and its application to default modelling. *Applied Mathematical Finance*. https://doi.org/10.1080/1350486X.2018.1481439.

Kaushansky, V. and Reisinger, C. (2018). Simulation of a simple particle system interacting through hitting times. *Discrete and Continuous Dynamical Systems*, To appear.

Kiesel, R. and Scherer, M. (2007). Dynamic credit portfolio modelling in structural models with jumps. *Preprint, Universität Ulm*.

Kim, I. J., Ramaswamy, K., and Sundaresan, S. (1993). Does default risk in coupons affect the valuation of corporate bonds?: A contingent claims model. *Financial Management*, 22(3):117–131.

Kluge, T. (2002). Pricing derivatives in stochastic volatility models using the finite difference method. *Diploma thesis, Technische Universiteta Chemnitz*.

Kolk, M. and Pedas, A. (2009). Numerical solution of Volterra integral equations with weakly singular kernels which may have a boundary singularity. *Mathematical Modelling and Analysis*, 14(1):79–89.

Kolk, M. and Pedas, A. (2013). Numerical solution of Volterra integral equations with singularities. *Frontiers of Mathematics in China*, 8(2):239–259.

Kolk, M., Pedas, A., and Vainikko, G. (2009). High-order methods for Volterra integral equations with general weak singularities. *Numerical Functional Analysis and Optimization*, 30(9-10):1002–1024.

Kythe, P. K. (2011). *Green's Functions and Linear Differential Equations: Theory, Applications, and Computation*. CRC Press.

Lando, D. (1998). On Cox processes and credit risky securities. *Review of Derivatives research*, 2(2-3):99–120.

Ledger, S. and Sojmark, A. (2018). At the mercy of the common noise: blow-ups in a conditional McKean–Vlasov problem. *arXiv preprint arXiv:1807.05126*.

Leland, H. E. (1994). Corporate debt value, bond covenants, and optimal capital structure. *The Journal of Finance*, 49(4):1213–1252.

Leland, H. E. and Toft, K. B. (1996). Optimal capital structure, endogenous bankruptcy, and the term structure of credit spreads. *The Journal of Finance*, 51(3):987–1019.

Li, D. X. (2000). On default correlation: A copula approach. *Journal of Fixed Income*, 9(4):43–54.

Linz, P. (1985). *Analytical and Numerical Methods for Volterra Equations*. SIAM.

Lipton, A. (2001). *Mathematical Methods For Foreign Exchange: A Financial Engineer's Approach*. World Scientific.

Lipton, A. (2002). Assets with jumps. *Risk*, 15(9):149–153.

Lipton, A. (2016). Modern monetary circuit theory, stability of interconnected banking network, and balance sheet optimization for individual banks. *International Journal of Theoretical and Applied Finance*, 19(6). doi: 10.1142/S0219024916500345.

Lipton, A. (2017). *Financial Engineering: Selected Works of Alexander Lipton*. World Scientific.

Lipton, A., Kaushansky, V., and Reisinger, C. (2018). Semi-analytical solution of a McKean-Vlasov equation with feedback through hitting a boundary. *arXiv preprint arXiv:1808.05311*.

Lipton, A., Pesavento, U., and Sotiropoulos, M. (2014). Trading strategies via book imbalance. *Risk, April*, pages 70–75.

Lipton, A. and Savescu, I. (2014). Pricing credit default swaps with bilateral value adjustments. *Quantitative Finance*, 14(1):171–188.

Lipton, A. and Sepp, A. (2009). Credit value adjustment for credit default swaps via the structural default model. *The Journal of Credit Risk*, 5(2):123–146.

Lipton, A. and Sepp, A. (2011). Filling the gaps. *Risk Magazine*, 24(10):78:66–71.

Lipton, A. and Sepp, A. (2013). Credit value adjustment in the extended structural default model. In Lipton, A. and Rennie, A., editors, *The Oxford Handbook of Credit Derivatives*, chapter 12, pages 406–463. Oxford University Press.

Lipton, A., Song, J. Z., and Lee, S. (2007). Systems and methods for modeling credit risks of publicly traded companies. US Patent 7,236,951.

Longstaff, F. A. and Schwartz, E. S. (1995). A simple approach to valuing risky fixed and floating rate debt. *The Journal of Finance*, 50(3):789–819.

Maleknejad, K., Hashemizadeh, E., and Ezzati, R. (2011). A new approach to the numerical solution of Volterra integral equations by using Bernstein's approximation. *Communications in Nonlinear Science and Numerical Simulation*, 16(2):647–655.

Maleknejad, K., Sohrabi, S., and Rostami, Y. (2007). Numerical solution of nonlinear Volterra integral equations of the second kind by using Chebyshev polynomials. *Applied Mathematics and Computation*, 188(1):123–128.

Marshall, A. W. and Olkin, I. (1967). A multivariate exponential distribution. *Journal of the American Statistical Association*, 62(317):30–44.

Merton, R. C. (1974). On the pricing of corporate debt: The risk structure of interest rates. *The Journal of Finance*, 29(2):449–470.

Nadtochiy, S. and Shkolnikov, M. (2017). Particle systems with singular interaction through hitting times: application in systemic risk modeling. *arXiv preprint arXiv:1705.00691*.

Nadtochiy, S. and Shkolnikov, M. (2018). Mean field systems on networks, with singular interaction through hitting times. *arXiv preprint arXiv:1807.02015*.

Nakatsukasa, Y., Noferini, V., and Townsend, A. (2015). Computing the common zeros of two bivariate functions via bézout resultants. *Numerische Mathematik*, 129(1):181–209.

Nielsen, L., Saá-Requejo, J., and Santa-Clara, P. (1993). Default risk and interest rate risk: The term structure of credit spreads. *INSEAD working paper*.

Noble, B. (1969). Instability when solving Volterra integral equations of the second kind by multistep methods. In *Conference on the Numerical Solution of Differential Equations*, pages 23–39. Springer.

Persson, P.-O. (2004). *Mesh generation for implicit geometries*. PhD thesis, MIT.

Pooley, D. M., Vetzal, K. R., and Forsyth, P. A. (2003). Convergence remedies for non-smooth payoffs in option pricing. *Journal of Computational Finance*, 6(4):25–40.

Reis, G. d., Smith, G., and Tankov, P. (2018). Importance sampling for McKean-Vlasov SDEs.

Reisinger, C. and Whitley, A. (2014). The impact of a natural time change on the convergence of the Crank–Nicolson scheme. *IMA Journal of Numerical Analysis*, 34(3):1156–1192.

Ricketson, L. (2015). A multilevel Monte Carlo method for a class of McKean–Vlasov processes. *arXiv preprint arXiv:1508.02299*.

Rubinstein, L. (1971). *The Stefan Problem*, volume 27 of *Translations of Mathematical Monographs*. American Mathematical Society, Providence, RI.

Sepp, A. (2004). Analytical pricing of double-barrier options under a double-exponential jump diffusion process: applications of Laplace transform. *International Journal of Theoretical and Applied Finance*, 7(02):151–175.

Sepp, A. (2006). Extended CreditGrades model with stochastic volatility and jumps. *Wilmott Magazine*, pages 50–62.

Spiliopoulos, K. (2015). Systemic risk and default clustering for large financial systems. In *Large Deviations and Asymptotic Methods in Finance*, pages 529–557. Springer.

Szpruch, L., Tan, S., and Tse, A. (2017). Iterative particle approximation for McKean–Vlasov SDEs with application to Multilevel Monte Carlo estimation. *arXiv preprint arXiv:1706.00907*.

Tikhonov, A. N. and Samarskii, A. A. (1963). *Equations of Mathematical Physics*. Dover Publications, New York. English translation.

Toivanen, J. et al. (2016). Application of operator splitting methods in finance. pages 541–575.

Trefethen, L. N. (2013). *Approximation Theory and Approximation Practice*. SIAM, Philadelphia.

Valuzis, M. (2008). On the probabilities of correlated defaults: a first passage time approach. *Nonlinear Analysis: Modelling and Control*, 13(1):117–133.

Zhou, C. (2001a). An analysis of default correlations and multiple defaults. *Review of Financial Studies*, 14(2):555–576.

Zhou, C. (2001b). The term structure of credit spreads with jump risk. *Journal of Banking & Finance*, 25(11):2015–2040.