Charles University in Prague

Faculty of Social Sciences

Institute of Economic Studies

MASTER THESIS

Monte Carlo simulation of Counterparty Credit Risk

Author: **Robert Havelka**

Supervisor: **PhDr. Boril Šopov, M.Sc., LL.M.**

Academic Year: **2014/2015**

Declaration of Authorship

1. Hereby I declare that I have compiled this master thesis independently, using only the listed literature and sources.
2. I declare that the thesis has not been used for obtaining another title.
3. I agree on making this thesis accessible for study and research purposes.

Prague, May 10, 2015

Signature

Acknowledgments

The author is grateful to PhDr. Boril Šopov, MSc., LL.M. for invaluable comments and help. My thanks also go to our Alma Mater, which offers an academic environment supporting creativity and achievement.

Abstract

The counterparty credit risk is particularly hard to simulate and this thesis is only the second work so far, which considers effective simulation of couterparty risk. There are two new approaches to stochastic modelling, which are useful with respect to efficient simulation of counterparty risk. These are Path-Dependent Simulation (PDS) and Direct-Jump to Simulation date (DJS). It had been show that DJS is far more effective, when it comes counterparty risk simulation of path-independent derivatives. We focus on a portfolio of interest rate swaps, which are effectively path-dependent. DJS approach yields estimates with much lower variance than PDS approach. But as expected, the DJS is also much more computationally intensive. The increase in computing time in majority of cases wipes out any gains in lower variance and PDS approach is shown to be more effective, when computing time is taken into account. We also show that in practice the convergence rate of Monte Carlo method significantly underestimates the true reduction in variance, which can be achieved with increasing number of scenarios.

JEL Classification	C02, C15, C63, G01, G12, G32
Keywords	Monte Carlo, CVA, Exposure, Variance

Author's e-mail	robberth.cz@gmail.com
Supervisor's e-mail	boril.sopov@gmail.com

Abstrakt

Kreditní riziko protistrany je obzvlášť těžké simulovat a naše práce je teprve druhá v pořadí, která se zabývá efektivní simulací kreditního rizika protistrany. Dva nové přístupy ke stochastickému modelování ve spojení s kreditním rizikem protistrany se objevili v podobě přístupů "Path-Dependent Simulation" (PDS) a Direct-Jump to Simulation date (DJS). V minulosti bylo ukázano, že DJS přístup je mnohem efektivnější než PDS pokud vezmeme v úvahu pouze deriváty závislé-od-cesty. My bereme v úvahu portfolio swapů s výměnou úrokové sazby, které jsou efektivně jako derivát závislé-od-cesty. DJS přístup vede k odhadům s mnohem větší přesností než přístup PDS. Jak se dalo i očekávat tak přístup DJS je mnohem náročnější na výpočetní sílu. Tato výpočetní náročnost ve většině případů avšak přesahuje jakékoli zisky přesnosti přístupu DJS. PDS přístup je proto ve většinu případů efektivnější

pokud jde o simulaci kreditního rizika protistrany. Také se ukázalo, že konvergenční poměr metody Monte Carlo může značně podhodnotit zisky přesnosti, kterých je možno dosáhnout v praxi při zvýšení počtu scenárií.

Klasifikace JEL	C02, C15, C63, G01, G12, G32
Klíčová slova	Monte Carlo, CVA, Expozice, Rozptyl

E-mail autora	`robberth.cz@gmail.com`
E-mail vedoucího práce	`boril.sopov@gmail.com`

Contents

List of Tables

List of Figures

Acronyms

DJS Direct-Jump to Simulation Sate
PDS Path-Dependent Simulation
CVA Credit Value Adjustment
OTC Over-The-Counter
BIS Bank for International Settlements
CDS Credit Default Swap
BM Brownian Motion
SDE Stochastic Differential Equation
EPE Expected Positive Exposure
DVA Debt Value Adjustment
FVA Funding Value Adjustment
TAF Term Auction Facility
ISDA International Swaps and Derivatives Association
EEPE Effective Expected Positive Exposure
PEE Peak Expected Exposure
NPV Net Present Value
MSE Mean Square Error
BCBS Basel Committee on Banking Supervision

Master Thesis Proposal

Institute of Economic Studies
Faculty of Social Sciences
Charles University in Prague

Author:	**Bc. Robert Havelka**	Supervisor:	PhDr. Boril Šopov, M.Sc., LL.M.
E-mail:	robberth.cz@gmail.com	E-mail:	boril.sopov@gmail.com
Phone:	733467962	Phone:	987654321
Specialization:	*Economics and Finance*	Defense Planned:	June 2015

Proposed Topic:

Monte Carlo simulation of Counterparty Credit Risk

Motivation:

Prevalence of losses during the last financial crisis can be ascribed to losses caused by Counterparty Credit Risk. Such risk faces a financial institution on OTC markets, where central authority is non-existent, and thus there is risk of default of the counterparty one is in the business with. There was an increase in academic contributions to this field of risk management after the financial crisis occurred, since Counterparty Credit Risk was very much not taken into account prior the crisis and the development of this field was inadequate as signified by such huge losses.

The most important measure of Counterparty Credit Risk is the Credit Value Adjustment (CVA), which is its market value and is the difference between risk-free portfolio and the value of portfolio accounting for the possible default of the counterparty.

Calculation of CVA alone is very computationally-intensive task, especially when an institution holds thousands of contracts. As a results, CVA has to be calculated less often than it would be needed. Thus there is need for research paper in order to establish appropriate and efficient framework for simulation of Counterparty Credit Risk, it will be the first goal of this thesis to achieve that. This efficient framework not only means picking the right mathematical models for simulations, but also establishing the right sequence of steps in data processing etc..

There are several approaches in order to decrease the complexity of CVA calculations, such as analytical approximations or variance-reduction techniques in case when the CVA is calculated using Monte Carlo (MC) simulations. The second goal will be to compare such approaches in terms of their preciseness and efficiency and we will also try apply variance-reduction techniques ourselves.

Hypotheses:

1. Hypothesis #1: Analytical approximation to CVA calculation leads to less precise estimates than in case of MC simulation.
2. Hypothesis #2: Importance sampling can be applied in order to reduce the complexity of MC simulation of CVA.
3. Hypothesis #3: Analytical approximations are relatively less useful for more complex derivatives.

Methodology:

We will carry out our practical simulations on different kinds of derivative financial instruments, they will differ by their complexity and path-dependence. These two criteria are important since the results will likely differ for path-dependent and non-path-dependent derivatives, that is one approach for CVA simulation will produce good results in one case, but unsuitable results for another. We will choose two path-dependent and two non-path dependent derivatives and in both categories we will choose one more complex derivative and one relatively simple in order to best approximate the reality of OTC markets. Likely the simulations for simple derivatives will not require as much preciseness as those for more complex derivatives, that is simple approximations will do

relatively worse for more complex derivatives. This could be taken as extension of one of our hypotheses.

Practical simulations alone will be based on all the possible approaches already developed in the literature, such as the simplest analytical approximations and the analytical approximations with more sophistication (Vrins, 2013). Simplest analytical approximations are for instance simple summation techniques, which do not compute any integrals and only roughly approximate them. Other important methods of simulation will include Monte Carlo simulations without variance-reduction and MC simulations with variance-reduction techniques included (Antonov et al., 2011, Brigo et al, 2013). MC approaches to the problem were developed in the literature, but these are very scarce so we will develop our own if possible. Our approach should be using the importance sampling, which have not yet been used.

These simulations should then uncover what methods are most appropriate for different derivatives. Our measures will be trade-offs between preciseness, speed and difficulty for implementation.

There is no academic article that compares different approaches to MC simulation of Counterparty Credit Risk and so we can't relate to other studies. Only available literature are general textbooks on statistics of MC simulations and academic articles, which present individual solutions to MC simulation of Counterparty Credit Risk, no aggregation of such solutions is found in the literature.

Expected Contribution:

The main contribution lies in comparing different approaches to simulation of Counterparty Credit Risk already developed in the literature in terms of their preciseness and speed. No academic paper carries out such task. Also we will try to find techniques most appropriate for different kinds of derivatives. After that we can establish if more sophisticated implementations will pay off to potential institutions interested in CVA simulation or if more simple approaches suffice. There is also need for establishment of efficient framework for CVA simulation.

Another contribution will be development of our own efficient Monte Carlo estimator based on importance sampling. Monte Carlo simulations find their applications more frequently now, if it is in the field of physics, computer science or finance. This is due to increased demand for numerical simulations, which often leads to better solutions of the problem than analytical formulation.

Outline:

1. Introduction: I will briefly introduce the topic of Counterparty Credit Risk, explain the importance of this topic for financial sector and summarize the nature of mathematical simulations in this area.
2. Counterparty Credit Risk: This part will delve into specifics of Counterparty Credit Risk, we will go through various definitions and theorems and discuss more advanced topics needed for the practical simulations.
3. Monte Carlo simulations and variance-reduction techniques: Here we will deal with the theoretical underpinnings of MC simulations, which has to understood in order to be used. Also topic of variance-reduction techniques will be introduced, but only in the form needed for the practical simulations.
4. Practical simulations: Here we will comment and take reader through our simulations, showing our results and provide instructions on how to replicate these results using Matlab.
5. Results: I will summarize my findings and their implications for practitioners in financial sector.
6. References/Bibliography.

Core Bibliography:

Antonov, A. and Issakov, S. and Mechkov, S. (2011). Algorithmic Exposure and CVA for Exotic Derivatives (November 17, 2011). Available at SSRN: http://ssrn.com/abstract=1960773 or http://dx.doi.org/10.2139/ssrn.1960773

Basel Committee on Banking Supervision (2010), "Basel III: A Global Regulatory Framework for More Resilient Banks and Banking Systems," http://www.bis.org/publ/bcbs189_dec2010.pdf, December 2010.

Brigo, D., Alfonsi, A. (2005): Credit Default Swaps and Option Pricing with SSRD Stochastic Intensity and Interest-Rate Model, Finance & Stochastics, Vol. IX(1), 2005.

Brigo, D., M. Morini and A. Pallavicini (2013). Counterparty Credit Risk, Collateral and Funding: With pricing cases for all asset classes, John Wiley & Sons Ltd.

Kenyon, C., Stamm, E., (2012), Discounting, Libor, CVA and Funding, Interest Rate and Credit Pricing, (Applied Quantitive Finance) Palgrave Macmillan.

Pykhtin, M. (2005). Counterparty Credit Risk Modelling: Risk Management, Pricing and Regulation, Risk Books, 2005.

Pykhtin, M., and S. Zhu (2007). A Guide to Modeling Counterparty Credit Risk, GARP, July/August 2007, issue 37, pp. 16-22.

Vrins, F. (2011). "Credit Risk - Getting CVA Up and Running." Risk : Managing Risk in the World's Financial Markets. 2011.

Author **Supervisor**

Chapter 1
Introduction

Counterparty credit risk came to be of great importance since the last financial crisis. This does not mean counterparty risk was not relevant for behavior of financial markets before that, it just was not recognized as such (Borio, 2004). It is estimated that two thirds of losses on OTC markets related to derivatives were caused in fact by counterparty credit risk, specifically due to declines of creditworthiness of counterparties in the market (BIS, 2011). These losses are mark-to-market losses and they can be represented as changes in the Credit Value Adjustment (CVA). That means majority of the losses were not caused by bankruptcies or defaults of these counterparties, but rather due to changing perceptions of market participants. The insufficient emphasis on counterparty credit risk before the financial crisis can be most easily deducted from the fact that Basel II does not address this topic at all, Basel II addresses only actual defaults and changes in credit rating with respect to capital rules that market participants have to fulfill (BCBS, 2010). Basel III recognizes counterparty credit risk in full extent and capital requirements specifically demand that market participants set aside capital that covers potential losses resulting from decline in creditworthiness of counterparties with whom they are in business with (BCBS, 2010).

Thus CVA losses can be seen as a problem to be tackled as it caused the majority of the losses during the last financial crisis and at the same time it was not given the degree of attention it should have received. CVA is the difference between the value of portfolio under the assumption of risk-neutrality and portfolio which takes into account the possibility of counter-party's default and changes in its creditworthiness (Pykhtin and Zhu, 2007). Hence CVA can be seen as a value that market assigns to counterparty credit risk. CVA is the central measure when it comes to pricing of counterparty credit risk (Brigo et al, 2013). This risk is not observed on usual

exchanges, where degree of uniformity is required and potential losses from counterparty credit risk are covered by the exchange alone, namely by the clearing house. The clearing house clears and settles the transactions and in case of default of the counterparty steps in. The step in means that clearing house settles the transaction for the failing counterparty and in the mean time also receives collateral from the counterparties to account for this risk.

The CVA alone is very hard to calculate. The computational intensity needed to quantify counterparty credit risk is significant due to number of reasons. This results from properties of simulations, which require multiple steps and there is number of complex steps that need to carried out before the estimates of CVA and other measures of counterparty credit risk can be obtained (Brigo et al, 2013). Because of multiple steps involved in these simulations it implies that an institution desiring to quantify counterparty risk has to have complex internal system, but that is not the case often (McKinsey, 2010). Tackling such problem with spreadsheet calculations is common and it is very unfit for such task as spreadsheets are too slow, too simplistic and without any possibility to build more complex model for quantification.

1.1 Why so complicated ?

If an institution participating in OTC markets assumes that its counterparties are not default-free, then it should desire the pricing of counterparty risk. The hindrance of this task is the sheer computational intensity, which results due to the need to extrapolate future. As we will find out CVA can be approximated by the following analytical formula

$$CVA = (1-R)E[E[1\{\tau \leq T\}V(\tau))|\tau]] = \int_0^T E[V(t)]dF(t)$$

$V(t)$ is the so-called credit exposure process and $F(t)$ is the default probability distribution for a given counterparty. The first step is to set a future horizon T, the set of risk factors as a drivers of the portfolio's components and the stochastic model, which describes the evolution of these risk factors. The previous formula can be approximated as

$$\widehat{CVA} = \sum_{i=1}^{n} E[V(t_i)]\Delta F(t_i)$$

With the particular choice of discretization scheme, that is the selection of valuation points $t_i, i = 1, ..., n$ and their closeness, we see that have to calculate another n expectations and the default probability distribution $F(t_i)$ in order to obtain estimate of the initial expectation. For every time bucket $[t_{i-1}, t_i], i = 1, ..., n$, where n can range from 15 to infinity, one has to generate at least m scenarios of risk factors and then valuate the portfolio for each of the risk factors in order to obtain a good estimate of $E[V(t_i)]$ for every time bucket. The range m assumes can be between 1000 and 4 billion (Brigo et al, 2013). The valuation alone can be intensive and we have to do that $n * m$ times at least, for variety of instruments. Carrying out all these steps is going to yield us a single estimate of the initial expectation.

However, with a single estimate of CVA we will not know how precise the estimate is. We would be required to repeat the previous steps many times to obtain a variance, a measure of estimate's accuracy.

The previous paragraphs only show the complexity of pricing the counterparty risk and not risk measurement. Risk measurement would require to obtain Value-at-risk of CVA. We would have to repeat the whole computation for every time-bucket in order to obtain CVA estimates for every time bucket and generate the distribution of CVA losses. This essentially leads to situation with many nested levels of simulations, sub-simulations and sub-sub-simulations (Brigo et al, 2013).

As we stressed in the beginning of this section, the formula for calculating CVA we present is analytical, in its entirety CVA should be calculated as a value of our portfolio given the default of counterparty, not use its default probability((Pykhtin and Zhu, 2007). This would include generating number of exposures and simulating the default process at the same time. Given that the default is highly unlikely, the number of exposures to be generated so that for some exposures the counterparty actually defaults is potentially huge. Not a sensible approach without utilizing importance sampling at least.

Another potential problems in counterparty risk include the amount of measures, which are to be calculated with connection to counterparty risk. Even for regulatory purposes, CVA is not the only quantity required.

1.2 The real world challenges

Financial institutions are investing large sums of money into developing systems for measuring counterparty credit risk, which would enable them to calculate required

capital to be set aside covering this risk (Ghamami and Zhang, 2013). These systems are based on Monte Carlo simulations. But even systems not based on spreadsheet calculations can not produce results financial institutions would like to. It is too computationally intensive to quantify counterparty credit risk in its entirety.

Basel III includes standardized methods for CVA calculation that banks has to apply if they do not opt to have their own internal system, which regulator has to permit ex ante (BCBS, 2010). These standardized methods rely on various restrictive assumptions and employ analytical approximations to calculate Value-at-risk of the change in CVA. However, assumptions that simplify the calculation of CVA may lead to estimates that do not approximate the true CVA very well (Brigo et al, 2013).

CVA has many features that are advanced and too complex to model. Such as the wrong way risk, which describes the situation, when there is dependence between the probability of default and size of the exposure (Pykhtin and Zhu, 2007). One often assumes that wrong way risk does not exist.

Model for CVA calculation also depends on the type of instrument in question (Antonov et al, 2011). Different approaches are taken to calculate CVA for credit instruments, such as Credit Default Swap (CDS), interest rate derivatives and commodity instruments (Brigo et al, 2013). Financial institutions on average hold thousands of instruments in their portfolio of varying type. As Brigo et al (2013) point out, it is almost impossible task to actually standardize CVA calculations across instruments and implement it in internal model systems.

1.3 Goals of this thesis

We want to compare different approaches for CVA calculation presently available and potentially develop some of our own. For the case of Monte Carlo method it is possible to employ so-called variance-reduction techniques, which should yield improvements in the speed of computations. Variance-reduction techniques try to exploit the specific or some general knowledge about the problem, that would increase the accuracy of true estimates for the same number of iterations. Also the nature of counterparty risk allows to use more effective versions of Monte Carlo simulations, where we fundamentally change the approaches taken so far in this area.

However, the results can be misleading, when employing these techniques within Monte Carlo method and different Monte Carlo methods. It is possible to increase the accuracy of the estimates by some factor, but at the same time to increase significantly

computational intensity per every iteration as to wipe out any gains in efficiency obtained so far. In fact we can increase the overall computational intensity when using some more accurate approaches. Thus, we will have to use appropriate measures to account for such differences. These measures most often mix bias, variance and computing time together in order to obtain a single quantity summarizing the three.

The portfolio of interest rate swaps will be our main interest. This is instrument for which CVA is calculated most often (Pykhtin and Zhu, 2007). Also interest rate swaps are relatively less demanding, when it comes to dependency modeling (Brigo and Alfonsi, 2005). We can take assumptions about wrong way risk and use a simpler framework for simulations. To our best knowledge, there is no academic article that take these tasks upon itself. Research in this area is very limited.

1.4 Structure

In chapter 2 we will start describing what Monte Carlo simulations are and how useful they are in derivative pricing. The next sections of the chapter will talk about the random number generation, sampling from probability distributions and also variance-reduction methods. Chapter 3 enlightens the topic of stochastic modeling and describes the processes and models we will later employ. This includes the generation of the Brownian motion, the description of models of the short-rate and making the important distinction between two approaches to stochastic simulation - Path-Dependent simulation (PDS) and Direct-Jump to simulation date (DJS). Finally, the chapter 4 utilizes the previous findings and applies it to the simulation of counterparty credit risk. We also specify what a default really is and what exposure means. Practical simulations are carried out in chapter 5. We conclude with the chapter 6.

Chapter 2
Monte Carlo simulations

Methods of Monte Carlo simulations involve a significant range of algorithms, which serve many purposes, but there are few generalization that can be made. First, Monte Carlo methods are methods, which are based on repeated random sampling and that is generally the first step of all algorithms in this domain. The first step serves as a way how to address various problems at hand. Random sampling consists of generating random numbers of huge quantities. Generated numbers should satisfy particular criteria. which can include dependence structure, size, other distributional properties. These criteria can be enforced using statistical tests, for instance to test for their independence.

The problem at hand can be either deterministic of include some form of uncertainty. Problems that involve some form of uncertainty are area, where Monte Carlo method is very suitable. This is often the case, that uncertainty can not be tackled otherwise and Monte Carlo is the only viable option. Analytical approximations can be used in many case for expression of the properties of uncertain phenomena and it can lead to relatively simple solution of the problem, but at the expense of precision and accuracy. On the other hand, employing Monte Carlo can be very computationally intensive, time-consuming and the implementation more complex and difficult. To summarize, Monte Carlo methods are used determination and understanding the properties of various phenomena using randomness.

The disciplines that use Monte Carlo methods extensively are Mathematics and Physics, as well as the area of Finance. Specifically, the problem of integration is tackled by Monte Carlo methods well even in cases when analytical approach does not yield suitable results. This is the most common application. Monte Carlo methods is especially suitable for evaluating multidimensional integrals as it has properties that set it apart from deterministic methods for computing multidimensional

integrals. This is crucial property as many problems in Finance are inherently multidimensional and representable as an integral. Monte Carlo methods involve also techniques that can speed up the evaluation of integrals increasing the preciseness of the estimates with the employment of so-called variance-reduction techniques. Variance-reduction techniques reduce the computational intensity needed for estimation of desired quantities by utilizing general or specific properties of the problem at hand. Analytical representations alone can be employed with variance-reduction techniques. This introduces degree of flexibility into Monte Carlo method others do not have.

2.1 Pricing derivatives

Derivative is kind of financial instrument, whose value depends on the behavior of other financial assets of more simple nature. Interest rate swap is one such derivative. Monte Carlo simulations are useful in derivative pricing due to the ability to state derivative prices as expectations. To be able to obtain such simplified expressions of derivative's price a few assumptions have to be made. Glasserman (2004) summarizes three most important principles of derivative pricing as follows:

1. The price of a derivative should be equal to the cost of replicating or emulating the derivative's payoff by employing other financial assets.
2. Such replication or emulation is possible in the so-called complete market and every derivative is replicable in this market. Perfect hedge is possible for every derivative contract. In complete market there is also unique probability measure under which discounted asset prices are martingales.
3. Finally, price of a derivative is an expectation of derivative's discounted payoff under the assumption of complete market and the resulting unique probability measure.

The unique probability measure is called the risk-neutral measure and its existence is necessary for market completeness. The complete market assumption is crucial as otherwise it is not ensured that a derivative can be perfectly replicated by other financial assets. One important assumption within market completeness is the assumption of no-arbitrage, which means that there is no risk-free profit to be made and it effectively assumes that all investments are risk-free and hence returns on all assets are equal to the risk-free rate of return. To discount an asset price and obtain a present value of the payoff we only use a risk-free rate. It would be possible to estimate price

of a derivative upon finding a specific replicating strategy for it and its cost. The first principle says exactly that, but such a strategy would be relatively more complicated. Instead we use the advantages of risk-neutral measure and its world.

To estimate the price of a derivative through an expectation, first one has to simulate the behavior of the underlying assets, these can be asset price or interest rate, through large number of scenarios and then quantify the derivative's payoff on each of them. Discount the payoffs to present by using a risk-free rate as a discount factor and average over all paths.

2.2 Monte Carlo Method in Finance

Important parts of finance consist of valuation, risk measurement and analysis of various financial market instruments, portfolios of instruments, investments. However, there is inherent uncertainty contained in these tasks and Monte Carlo method can account for such uncertainty as mentioned. Uncertainty in finance can be expressed for instance by future path of price of some underlying asset, which can influence the price of derivative security, or the future development of interest rates, the level of risk. Because the development of these variables is uncertain with respect to future, we can not pinpoint or guess one specific path of variable's development or one specific price path of an underlying asset that would eventually turned out to be true. Every price path may be interpreted as a single realization of the future. Ideally we should be able to account for all possible realizations of the future and weight them according to the probability of their occurrence. Monte Carlo method accounts for this uncertainty by simulating many price paths, that is simulating a process, which underlies the problem to be solved. Simulated paths are then used as an input to a model, which produces the estimate of the problem at hand.

In case of an Asian option, where payoff is determined as an average of price of an underlying asset over a specified period, we would first take average of a single price path, it would yield the price of an option in one version of the future. To obtain the price of the option accounting for uncertainty, we would then take average over all versions of the future, represented by all price paths generated. The problems in finance, which do employ Monte Carlo method need to be specified as an integral. This comes as natural for many problems.

2.3 Basic principles of Monte Carlo method

As we have discussed, Monte Carlo method is suitable for evaluating integrals. The basic reason for that is the convergence rate of $\frac{1}{\sqrt{n}}$ for integral evaluation, which Monte Carlo offers (Glasserman, 2004). This means in order to decrease variance by two, we have to increase the number of scenarios by a factor of four. This convergence rate also applies for integrals of any dimension.

Another popular procedure for integral evaluation is the so-called trapezoidal rule for instance, which on the other hand has a convergence rate of $\frac{1}{n^{\frac{2}{D}}}$, where D is the number of dimensions our integral possesses (Glasserman, 2004). It is obvious that as the number of dimensions increases, the trapezoidal approach will become less and less efficient. The degradation in performance is too large to consider this technique in evaluating multidimensional integrals.

2.3.1 Statistics of Monte Carlo

In this section we will show how the convergence rate of Monte Carlo method is determined, while following closely Glasserman (2004). Suppose we have N i.i.d. realizations $C_1, ..., C_N$ of a given integral and suppose $\bar{C}_N = \frac{\Sigma_{i=1}^{N} C_i}{2}$, where $E[C_i] = C$, $Var[C_i] = \sigma_C^2 < \infty$. $\bar{C}_N$ is then the estimate of the integral with standard error of σ_C and C is the supposed true value. The central limit theorem would give us that

$$\frac{\bar{C}_N - C}{\frac{\sigma_C}{\sqrt{N}}} \rightarrow N(0,1); N \rightarrow \infty$$

which can be transformed as $\bar{C}_N - C \rightarrow N(0, \sigma_C^2)$ (Glasserman, 2004). And it is possible to replace the standard error with the standard sample deviation.

2.4 Generating Random Numbers

Generating random numbers is generally the first step of any Monte Carlo simulation, these numbers are then driving force behind all following computations (Glasserman, 2004). Thus it is important to describe and understand how random numbers are generated. Following sections will discuss methods for random number generation and show that these methods are all deterministic, no random elements are employed.

It begs to ask then if such numbers are truly random, when they are in fact generated by deterministic algorithms only. This question will be answered as well.

2.4.1 Random number generation in simple manner

We only have to restrict our attention to generating a sample from uniform distribution, i.e. $U[0,1]$, since such sample can be transformed into sample of arbitrary distribution. This is a question of efficiency as one would have to create new generator for every distribution individually, if it wouldn't be for this solution. Sample from $U[0,1]$ is in fact also the distribution of function values of any possible distribution as every probability distribution function is projected into the space of $[0,1]$. Follows the formal definition for generator of random numbers.

Definition 1. Random number generator is a mechanism that produces random variables $U_1, U_2, ...$, which satisfy following two properties:

1. $\forall i = 1,2,...$ we have that $U_i \in U[0,1]$, i.e. it is uniformly distributed on the interval $U[0,1]$
2. every pair of random variables U_i, U_j, where $i \neq j$, is independent of each other (Glasserman, 2004)

Generator that satisfies such properties should produce sequence of numbers with apparently no pattern set in them and the number of values in any given sub-interval should should be equal to the length of that particular sub-interval, assuming that the number of generated values is large enough. Follows the definition of our first simple generator of random numbers.

Definition 2. Linear congruential operator is recursive algorithm defined as follows:

$$x_{i+1} = ax_i \, mod \, k$$

$$u_{i+1} = x_{i+1}/k$$

where a is the multiplier, initial value of x_0 is the seed and $U = (u_1, u_2, ...)$ follows $U[0,1]$.

This operator is the simplest generator of random numbers. It uses deterministic means and is recursive in nature. It is effective in the sense that modulo operation can be very effectively implemented at lower level of computation in terms of handling the individual bits of number and then we only have to proceed with multiplication and division of two numbers, which is very fast as well (Glasserman, 2004).

2.4.2 Simple random number generation - dependency and length

However, any sequence of values

$$x_0, x_1, ..., x_n, n \in R$$

generated by this algorithms will always consists of sequences that repeat itself. Assume $k = 9, a = 5, x_0 = 2$, then the sequence of values generated using this operator looks like

$$1, 5, 7, 8, 4, 2, 1, 5, 7, 8, 4, 2, 1, 5, ...$$

This way we have achieved to produce a unique sequence of the length 6, which eventually repeats itself as observed. The number of values which linear congruential operator produces before repeating is called the period and the longest possible period is the full period, that is period with length of $k-1$ (Glasserman, 2004). This implies that k should be very large in order to be able to generate large number of random values, but also constants should be chosen carefully. The previous example shows the length of the unique sequence is not necessarily always equal to $k-1$.

Being able to generate higher number of independent random numbers is useful, as the distribution of these numbers will follow the uniform distribution more closely and we can also better satisfy the needs for generating long sequences of independent random numbers. This is often the case and the value of k can be regularly around 2^{32}. Not every operator produces sequence with full period and it depends on the parameters selected as it is evident from our example.

2.4.3 A seed

Another important property of our simple generator concerns the choice of so-called seed, i.e. x_0, which allows for reproducibility of our sequence of random values generated. If seed is given the same and parameters of the operator are kept unchanged, we will always receive sequence with the same numbers. This property can come in handy, if we want to retrace our steps during an individual simulation for instance.

Splitting random number sequences is also possible with linear congruential operators. Such property is useful when running our algorithm on parallel computing unit (Glasserman, 2004). We can simply split the sequence so that parallel computations are independent of each other. This is can be done by choosing different seeds,

but not chosen randomly. We pick two numbers from the sequence generated with particular seed, while keeping in mind that these two numbers has to be far away from each other in the sequence. To get such number we pick the first one and then skip ahead by l steps, i.e.

$$x_{i+l} = a^l x_i \; mod \, k$$

Then we can use the seeds x_0 and x_{l+k} to generate the sequence of random numbers, which have l spaces between themselves.

2.4.4 Advanced Generators

Linear congruential operators against all its simplicity are still useful, when it comes to advanced methods of random number generation. The next step are so-called Combined Generators, which as name suggest combine number of linear congruential operators together. It appears that such configuration leads to better dependence structure, which means higher degree of randomness in generated numbers and longer periods, while inheriting the favorable property of simplicity (Glasserman, 2004).

Combining linear congruential operators proceeds in following manner: Assume we have L such generators, $l = 1, 2, .., L$, each with parameters a_l, k_l. One approach is then to generate $u_{i+1,l}$ for every l and sum them. To obtain u_{i+1} we simply subtract from the sum its integer part. Such combined generator produces random sequences, which are shown to have no lattice structure and is able to generate random sequences with length far longer than the ones generated by individual generators (Glasserman, 2004).

2.4.5 Independence of Random Numbers

There are a few approaches for evaluating our random generators and the values they produce. Approaches include the lattice structure of generated values, spectral tests and discrepancy measures.

Lattice structure relies on the fact that if random variables $U_1, U_2, ..., U_n$ are independently and identically distributed and come from $U[0,1]$, then for any $d \leq n$ the random variables $U_1, ..., U_d$ are uniformly distributed in the d-dimensional space of $[0,1]^d$ (Glasserman, 2004). Thus to evaluate the linear congruential operator we can generate a sequence of random numbers and form points in 2-dimensional space using consecutive values in the sequence. If we follow this plan and plot the values

generated using linear congruential operator, we would see there is clear pattern and the numbers are hardly independent (Glasserman, 2004).

The graphs also can help us to pick the right parameter for our recurrence. Lattice structures are in turn analyzed by spectral tests, which try to quantify the degree of regularity in structure showed above. There are also other statistical tests that can be carried out, these include tests of independence or weak dependence, which are frequent in time series econometrics.

2.5 Sampling Methods

The previous section assures us of being able to generate a sequence of random numbers from $U[0,1]$. Now we approach the general topic of sampling from distributions we actually want to study, while utilizing the sample from $U[0,1]$. We will discuss two methods in use, which are most popular and that is the Inverse Transformation and Acceptance-Rejection method. Most of simulations in stochastic finance involves sampling from a few fundamental distributions, such as Normal distribution.

2.5.1 The Inverse Transform method

Assume we have a continuous cumulative distribution function F defined on some set X, i.e. $P(X \leq x) = F(x), \forall x \in X$, and we want to generate a sample from this distribution. Then if function F has its inverse well-defined, we can write

$$F(X) = U, U \sim U[0,1], X = F^{-1}(U)$$

. This is justified by following

$$P(X \leq x) = P(F^{-1}(U) \leq x) = P(U \leq F(x)) = F(x)$$

The last equality follows from the fact that any cumulative distribution function $F(x)$ is uniformly distributed between 0 and 1. The exact number between 0 and 1 signifies the percentile of the cumulative distribution function. The inverse is defined for all strictly increasing functions, while for non-decreasing functions the inverse still exists if precautions are taken (Glasserman, 2004). If a function is non-decreasing, there can be a single function value assigned to more than one value from the set X, which corresponds to a flat section. The cases, where the inverse of a function does

not exist in a closed form, techniques of numerical integration can be used, then our problem would reduce to solving equation of the form $F(x) = u$ for x.

2.5.1.1 The Inverse Transform method for discrete distributions

Discrete distributions can be handled without any need for explicit calculation of the inverse of cumulative distribution function (Glasserman, 2004). Assume we have a discrete random variable and its sample consisting of values

$$\{x_1, x_2, ..., x_n; n > 0, n \in R\}$$

and their probabilities

$$\{p_1, p_2, ..., p_n; n > 0, n \in R\}$$

Vector of values is sorted in ascending order. The cumulative distribution function of such variable is then given as $F(x_i) = \sum_{j=1}^{i} p_j$. To generate a sample X from this distribution we generate a sample from $U[0,1]$, i.e. $u_i, i = 1, 2, ..., n$, and find such pair of x_i, x_{i+1} for which

$$F(x_i) < u_i \leq F(x_{i+1})$$

The desired value of X is then x_{i+1}. The computational intensity of sampling from discrete distribution is directly observable. Apart from generating a sample from $U[0,1]$, we only have look for a certain value in a sorted array, which can be handled by binary search in logarithmic, i.e. in $O(log(n))$.

To apply this method we need to have ready the inverse of desired distribution we want to sample from. Most comfortable is to have an analytical formulation for the inverse function, this is not possible in many cases. The next method offers a remedy.

2.5.2 The Rejection Sampling

Another popular method for generating random samples from probability distributions is the so-called Rejection Sampling or Acceptance-Rejection method. Using this method, we do not generate a random sample from the desired distribution directly, instead we use a more viable distribution and generate values from this distribution, some of which get randomly rejected. The random rejection part of this method makes sure the generated sample is in fact distributed according to desired distribution even though we are sampling from different distribution.

We assume here that we face certain constraints in sampling from the desired distribution. Possibly, we do not have available the inverse function of the distribution of interest and can not use the the inverse transform method consequently.

Suppose that we have two probability densities f and g defined on some set X. The density f is the desired density, while the latter is a density from which random sample is readily to be drawn. The densities are related as such $f(x) \leq cg(x), c \in R, \forall x \in R$. The following algorithm is due to Glasserman (2004) and allow us to sample from f.

1. generate a sample x from density g and u from $U[0,1]$
2. accept x as sample from f, if $u \leq \frac{f(x)}{cg(x)}$
3. repeat steps 1,2 as necessary

Every sample x from density g is accepted with probability $\frac{f(x)}{cg(x)}$. Sample from Uniform distribution serves only as a metric for acceptance or rejection. Glasserman (2004) shows formally that this method is in fact correct, we omit the proof of that.

2.6 Variance Reduction methods

Monte Carlo simulations tend to be computationally intensive and consequently variance reduction methods help to develop modified Monte Carlo estimators with higher accuracy than crude versions of these estimators. The logic behind ambitions of this method is that in every estimator there is inherent error present and it is embodied by the variance. Variance is the error of estimates of given quantity. Given a constant number of simulation runs, the variance of crude MC estimator shall be higher than the variance of estimator with successfully implemented reduction of variance. Effectively, with such an estimator lower number of simulation runs is needed to achieve a given constant accuracy.

Application of these methods is not obvious and requires as good knowledge as possible about the model and problems with which the simulations concern itselves. To exploit its specific properties and not “generic application of generic methods” (Glasserman, 2004) one shall. The methods to be discussed are Control Variates and Stratified Sampling.

2.6.1 Control Variates

One of the straightforward variance-reduction methods to apply is Control Variates. Together with the variable of interest we generate as well some other variable for which we know the estimate errors. Then we use it to reduce the variance of variable of interest. The other variable is likely a one for which we already know how to handily estimate the expectation. As we will see later the potential gain is governed by the strength of the dependency between these two variables, the greater is the dependency, no matter if positive or negative, the more is to gain.

Assume we have two random variables (X,C) and i.i.d. sample of size n from both distributions. The goal is to estimate $E[X]$. This could be done by employing a sample mean $\bar{X}$, which is an unbiased estimate of $E[X]$. However, suppose we transform the draws from the distribution of X followingly

$$\hat{X}_i(\lambda) = X_i - \lambda(C_i - E[C]).$$

We assume a constant λ and the ability to easily obtain $E[C]$. To obtain $E[\hat{X}]$ we simply take a sample mean again as

$$\bar{X}(\lambda) = \frac{\sum_{i=1}^{n}(X_i - \lambda(C_i - E[C])}{n}$$

It is an average over all n transformed draws from X given a constant λ.

2.6.1.1 Unbiasedness

To prove the unbiasedness of the Control Variates estimator we write

$$E[\bar{X}(\lambda)] = E[\bar{X} - \lambda(\bar{C} - E[C])] = E[X] - \lambda E[C - E[C]] = E[X]$$

and we use the independence of the two samples and unbiasedness of sample means.

2.6.1.2 Variance

The variance of the estimator will however depend on the variance of control variable C, the correlation between X and C, and λ:

$$Var[\hat{X}_i(\lambda)] \quad = \quad Var[X_i - \lambda(C_i - E[C])] =$$

$$\begin{aligned} &= (E[X_i - \lambda(C_i - E[C])])^2 - E[X_i - \lambda(C_i - E[C])]^2 \\ &= (E[X_i])^2 - 2\lambda E[X_i(C_i - E[C])] + \lambda^2 E[C_i - E[C]]^2 \\ &= \sigma_X^2 - 2\lambda\sigma_{XC} + \lambda^2\sigma_C^2 \end{aligned}$$

where $Var[X,C] = \sigma_{XC}$, $Var[X] = \sigma_X^2$ and $Var[C] = \sigma_C^2$.

The crude estimator would exhibit variance of σ_X^2 and so in order for Control Variates estimator to reduce the variance the following inequality shall hold

$$2\lambda\sigma_{XC} > \lambda^2\sigma_C^2$$

The optimal value of λ for which the variance of our estimator is minimized is obtained by minimizing with respect to λ

$$\lambda^* = \frac{\sigma_{CX}}{\sigma_C^2}$$

2.6.1.3 Efficiency of Control Variate estimator

We need to determine what exactly governs the gains in lower variance, when successfully employing the control variate estimator. Also it would be advantageous to see how the magnitude of the gain is determined. We compare the variance of Control Variates estimator with variance of X as follows

$$\frac{Var[X - \lambda^*(C - E[C])]}{Var[X]} = 1 - \rho_{XC}^2$$

where ρ_{XC} is the linear correlation between X and C. The previous expression gives us the explanation about why successful application of Control Variates requires the variables to be highly correlated, no matter if positively or negatively.

2.6.1.4 Challenges of application

One of the challenges of applying this estimator is the additional work needed for generating sample from the distribution of control variable and transforming the draws, hence rather high correlation is required for significant gains to be experienced. Also one does not know immediately the optimal value of λ. It is again possible to estimate it from the population using

$$\lambda^* = \frac{\sum_{i=1}^n (X_i - \bar{X})(C_i - \bar{C})}{\sum_{i=1}^n (C_i - \bar{C})^2}$$

As Glasserman (2004) points the estimate is equal to the slope in single least squares regression obtainable by regressing the control variates on the variable of interest. Such interpretation is interesting in case of multiple variates.

It is often the case that the control variable is some function of the variable of interest and additional computational demands are minor. Such situation can arise, when working under the risk-neutral measure and based on some assumptions we work out the expected value. For instance under the assumption of no-arbitrage, we know that discounted asset prices are martingales and their future value at any point in time is equal to the initial value. Hence initial value is then used as control variate (Glasserman, 2004).

2.6.1.5 Statistical properties of Control Variates estimators

The control variate estimator $\bar{X}(\lambda)$ is simple mean of its i.i.d. replications $X_i(\lambda), i = 1,..,n$ for given λ. To estimate the standard deviation $\sigma(\lambda)$ we can use the sample standard deviation defined as

$$S(\lambda) = \sqrt{\frac{\sum_{i=1}^{n}(X_i(\lambda) - X(\bar{\lambda}))^2}{n-1}}$$

Consequently, we can form the $(1-\alpha)$ confidence interval for our estimator as $(\bar{X}(\lambda) \pm z_{\alpha/2}\frac{\sigma(\lambda)}{\sqrt{n}})$, where $\sigma(\lambda)$ can be replaced by the sample standard deviation to keep confidence intervals valid asymptotically and $z_{\alpha/2}$ is the $(1-\alpha/2)$ quantile from standard normal distribution (Glasserman, 2004).

2.7 Stratified Sampling

Stratified sampling can be defined as a method, under which one systematically generates samples of desired size from disjoint subsets of the given sample space in order to estimate quantities of interest more efficiently and without a bias (Glasserman, 2004).

Suppose we have a random variable X defined on the real line. Then we partition the sample space into K disjoint subsets $A_1, \ldots, A_K$ so that we have $P(X \in \cup_{i=1}^{K} A_i) = 1$ and we want to find $E[X]$. We can write

$$E[X] = \sum_{i=1}^{K} E[X \cap X \in A_i] = \sum_{i=1}^{K} P(X \in A_i)E[X|X \in A_i] = \sum_{i=1}^{K} p_i E[X|X \in A_i]$$

If using general sampling method we generate enough draws from the given distribution, we would observe the number of draws from each subset A_i are converging to $\frac{p_i}{n}$, where n is the total number of draws (Glasserman, 2004). Stratified sampling gives more control over generating values from individual subsets. These subsets are named strata.

Assume we use a proportional sampling, so that if n is the total number of draws, we would draw exactly $n_i = np_i, i = 1,...,n$ observations from each of the strata A_i. Then if we obtain such stratified sample, to estimate $E[X]$ we use a sample mean this way

$$\bar{X} = \frac{\sum_{i=1}^{K} \sum_{j=1}^{n_i} X_{ij}}{n}$$

If we compare this to sample mean obtained the usual way, we find out stratified sampling eliminates all variation across strata, but not within them (Glasserman, 2004).

2.7.0.6 Generalized stratification

Generally, one does not restrict himself from generating sample from a variable of interest X, conditional on $X \in A_i$. But rather he samples conditional on $S \in A_i$, where S is another random variable called the stratification variable (Glasserman, 2004).

This is important generalization, as X can be often a function of S, where S are asset prices and X embodies the price of a derivative for instance. Since the price of a derivative is fully determined by the price of underlying asset, it makes more sense to use that asset price as a stratification variable, so that we can better explore the sample space. This way we determine the price of a derivative more efficiently.

In credit risk , the stratification variable is often the default time of obligor or counterparty. This stratification will allow to sample outcomes, which are relatively more important to the modeler. These generally happen less proportionately then the outcomes that involve the actual default (Pykhtin and Zhu, 2007).

Formally, if again $P(X \in \cup_{i=1}^{K} A_i) = 1$, then

$$E[X] = \sum_{i=1}^{K} P(S \in A_i) E[X|S \in A_i] = \sum_{i=1}^{K} p_i E[X|S \in A_i]$$

The above expression gives explicit instruction on how to apply stratified sampling in practice.

.

Chapter 3
Stochastic Modelling in Financial Engineering

In this chapter we will build the stochastic models to be used and study stochastic processes, which will be required for carrying out the practical simulations, including construction of random walk and dealing with the Brownian motion

3.1 Brownian motion

We start with the definition of Standard Brownian motion.

Definition 3. Standard Brownian motion is defined as a stochastic process $W(t)$, where $0 \leq t \leq T, T \in R$, that satisfies the following properties

1. $W(0) = 0$.
2. $t \to W(t)$ mapping is continuous function for $0 \leq t \leq T$.
3. The increments of Brownian motion $\{W(t_k) - W(t_{k-1}), 0 \leq t_k \leq T\}$ are independent.
4. $W(t) - W(s) \sim N(0, t-s)$ for any $0 \leq s < t \leq T$.
5. Consequently, we have $W(t) \sim N(0,t)$.

The previous definition is due to Glasserman (2004). The Brownian motion, whose definition is presented next, is generalization of the initial Standard Brownian motion.

Definition 4. The Brownian motion with drift parameter μ and the diffusion parameter σ^2, i.e. $BM(\mu, \sigma^2)$ and both parameters are non-negative a higher than zero

constants, is defined as a stochastic process $X(t)$ that satisfies the following transformation property:

$$\frac{X(t)-\mu t}{\sigma} \sim W(t)$$

where $W(t)$ is Standard Brownian motion.

The transformation property presented in the above definition leads to following stochastic differential equation (SDE)

$$dX(t) = \mu dt + \sigma dW(t)$$

and also can acquire form of

$$dX(t) = \mu(t)dt + \sigma(t)dW(t)$$

where drift and diffusion parameters are allowed to vary with time (Glasserman, 2004). Integrating the second equation we obtain

$$X(t) = X(0) + \int_0^t \mu(s)ds + \int_0^t \sigma(s)dW(s)$$

From the definition of Brownian motion we also have that $X(t) \sim N(\mu t, \sigma^2 t)$, if $X(t)$ is $BM(\mu, \sigma^2)$.

To construct a random walk using Standard Brownian motion, we assume a fixed set of dates $0 = t_0 < t_1 < ... < t_n = T$ and use the following recurrence

$$W(t_i) = W(t_{i-1}) + \sqrt{t_i - t_{i-1}}Z_i, i = 1, ..., n$$

where $Z = (Z_1, ..., Z_n) \sim N(0,1)$ are i.i.d. random variables (Glasserman, 2004). That will gives us the required vector $(W(t_0), ..., W(t_n))$ representing the random walk. To generate random walk using Brownian motion we have the following recurrence

$$X(t_i) = X(t_{i-1}) + \mu(t_i - t_{i-1}) + \sigma\sqrt{t_i - t_{i-1}}Z_i, i = 1, ..., n$$

Obviously, the generated values are not independent, since the next value in recurrence is always made conditional on the previous one and hence there is dependence structure present. This fact is important relatively to the methods of sampling we

will use in later parts of this thesis. The simulated vector

$$(W(t_0),...,W(t_n))$$

has the correct joint distribution of Standard Brownian motion at dates $t_1,...,t_n$, if simulated according to a given recurrence (Glasserman, 2004). It is clear that we can restrict ourselves to dealing with Standard Brownian motion only, since the Brownian motion is just a transformation of the first one and Standard Brownian motion is far simpler for demonstrating purposes.

3.1.1 Generating Brownian motion from multivariate distribution

Generating $(W(t_0),...,W(t_n))$ can be also reduced to generating a sample from multivariate normal distribution with appropriate dependence structure (Glasserman, 2004). To generate such a sample from multivariate normal distribution we need to obtain the covariance matrix and vector of means of the random walk. Vector of means is just a vector of zeros, since $E(W(t_i)) = 0, i = 0,..,n$. The covariance matrix is obtained again by considering the third property of Standard Brownian motion of independent increments. We have that

$$Cov(W(s),W(t)) = Cov(W(s),W(s) + (W(t) - W(s)) = min(s,t)$$

where $0 < s,t < T$. The covariance matrix is then represented as

$$Cov_{i,j} = min(t_i,t_j)$$

.

It is important to see that the above recurrence for constructing a random walk is significantly more efficient than generating a sample from multivariate normal distribution.

The latter is done usually with Cholesky decomposition next, where the covariance matrix is decomposed as $C = AA^T$, so that A is lower triangular matrix for which applies $A_{jj} = A_{ij}, \forall i \leq j$. To generate sample from $N(0,C)$ we do a multiplication AZ, where $Z \sim N(0,I)$. The calculation of the previous product is inefficient in the sense that it has many identical terms, which will be multiplied over and over again. The complexity of this approach is then $O(n^2)$ (Glasserman, 2004).

The random walk recurrence on the other hand never does the same work twice and produces complexity of $O(n)$. The quadratic complexity in the previous case cannot be reduced by an optimization technique, due to identical terms always being multiplied by different vectors of Z (Glasserman, 2004). Hence we will not consider drawing samples from multivariate distribution and rather use the more efficient recurrence.

3.2 Interest rate swap

The interest rate swap is defined as a derivative contract under which counterparties are to exchange two streams of payments, which are derived from a common notional principal. This principal, N, is never exchanged during the life of the derivative contract and its only use is to calculate the payment streams (Kenyon and Stamm, 2012). For standard interest rate swap, the first payment stream is based on a fixed rate K, and the other is based on a floating rate L.

Assume a discrete time grid $0 = t_1 < ... < t_n = T$, and interest rate swap with reset dates identical to dates $t_i, i = 1, ..., n$. Now, the floating rate for period $[t_{i-1}, t_i]$ is reset at t_{i-1} and then applied for the whole interval $[t_{i-1}, t_i]$, we denote this rate as $L(t_{i-1})$. The first payment exchange is done at t_1 and the last at t_n. The value of the fixed payment stream is as follows

$$\sum_{i=1}^{n} D(0,t_i)\beta_i NK$$

where β_i is the fraction of the year corresponding to the interval $[t_{i-1}, t_i]$, $D(0,t_i)$ is the discount factor. The value of the floating payment stream is similarly

$$\sum_{i=1}^{n} D(0,t_i)\beta_i NL(t_i)$$

The value of the swap to a counterparty paying fixed rate is then

$$\sum_{i=1}^{n} D(0,t_i)\beta_i N(L(t_i) - K)$$

Value of the swap to other counterparty is the previous payoff simply multiplied by -1. When these two payment streams are thought of as a two kinds of bonds, floating stream is just a floating rate bond easily replicable by investing N at t_0 at rate of $L(t_1)$ and at t_1 canceling the investment, paying the interest and reinvesting again

the principal for another period (Glasserman, 2014). By repeating this strategy the floating bond can be replicated with an initial investment of size N. Hence the value of floating payment stream is just N and we consequently obtain a value for fixed rate payer as

$$ND(0,t_n) + N - \sum_{i=1}^{n} D(0,t_i)\beta_i NK$$

At time t_0, when the contract is agreed upon, the fixed rate is set so that the present value of both payment streams are equal. If not, then naturally one of the counterparties would not engage in this contract due to facing a contract of negative value. The fair fixed rate K, is obtained from the previous expression as

$$K = \frac{1 - D(0,t_n)}{\sum_{i=1}^{n} D(0,t_i)\beta_i}$$

The previous derivation applies not only for swaps starting at t_0, but for any swap starting at any future time t_f. The fair forward swap rate starting at future time t_m and discounted to present is derived to be

$$S_m(t) = \frac{D(t,t) - D(t,t_{m+1})}{\sum_{i=1}^{n} D(t,t_i)\beta_i}$$

(Glasserman, 2004).

3.3 Stochastic interest rate model

The underlying variable affecting the price of interest rate swaps is the interest rate. Hence we need to choose a framework for modeling such a variable. To model the interest rate one picks the stochastic model of the short-rate, denoted as $r(t)$. Short rate is typically instantaneous and continuously compounded and generally it is assumed that it is normally distributed. Short rate approximates the behavior of interest rates. The last assumption of normally distributed short-rate leads to Gaussian short-rate models, which are the most popular and advantageous, when computational intensity is considered. The assumption of normality allows for analytical tractability, by yielding analytical pricing formulas for zero-coupon bonds and for the short-rate (Brigo and Mercurio, 2006).

The most important of all models, from which all models of the short-rate are derived, is the Vasicek model. We will its particular extension developed by Hull and

White (1990), which offers better properties and is still used in the risk management departments (Brigo and Mercurio, 2006).

Starting with Vasicek model, the equation determining the short-rate can be specified as

$$dr(t) = \alpha(\theta - r(t))dt + \sigma dW(t)$$

where $W(t)$ is again the Standard Brownian motion, α and θ are positive constants. The previous specification is due to the Glasserman (2004). This family of models is also called the mean-reversion family of models. Since $r(t)$ always converges to a constant θ. We see that if $r(t) < \theta$ the drift becomes positive and negative if $r(t) > \theta$. The constant α determines the speed of convergence of $r(t)$ to θ. The Hull-White extension rests in making the θ time-varying. Hence, the Hull-White model is then specified as

$$dr(t) = (\theta(t) - \alpha r(t))dt + \sigma dW(t)$$

with the drift function $\theta(t)$ specified as

$$\theta(t) = F_i(0,t) + \alpha F(0,t) + \frac{\sigma^2}{2\alpha}(1 - e^{-2\alpha t})$$

The previous specification of the drift function is due to MathWorks (2014).

Chosen interest rate model should be able to model interest rates well enough and provide good trade-off between computational tractability and realistic behavior of interest rates. The Vasicek model does not provide as good fit to an initial interest rate term structure, but the Hull-White extension should yield a much better fit (Brigo and Mercurio, 2006). The constant α could be made time-varying as well, but it leads to increased complexity, which need not provide as much benefits (Brigo and Mercurio, 2006). The constant σ specifies the volatility of the interest rate term structure. The main equation can be integrated and interest reader can see detailed results in Brigo and Mercurio (2006) and Glasserman (2004). Both of the references contain detailed derivations.

3.3.1 Generating the short-rate using Euler approximation

The algorithms for generating the short-rate using Euler approximation is given by Glasserman (2004) as

$$r(t_i) = r(t_{i-1}) + (\theta(t_{i-1}) - \alpha r(t_{i-1}))(t_i - t_{i-1}) + \sigma\sqrt{t_i - t_{i-1}}Z_i, Z \sim N(0,1), i = 1,..,n$$

Again the previous algorithm is of sequential nature, where short-rate for a given period is determined using a value of the short-rate in the previous period. Such an approach is always viable for PDS approach and in this case for DJS approach as well.

Due to the nature of stochastic discounting, which is path-dependent effectively, there is no point in reformulating the previous equation to yield an algorithm for always generating the short-rate given its initial value. It would be useless, since we are striving to make exposures at two given dates independent. As we will find out later, we always have to generate the short-rate path from the beginning and do not carry out any exposure generation until the last value of the short-rate path. If the short-rate used for generating these exposures has zero covariance terms along its path, it is not going to decrease any variance.

3.3.2 Zero-rate curve and the discount curve

We are interested first and foremost in the zero-rate curve for every given valuation date. If we denote $R(t,T)$ as the zero-rate at given date t for a tenor of up to $t-T$, we see that to obtain a yield curve at a given date t we have to generate function $T \rightarrow R(t,T)$. Such a yield curve will be generated at every simulation date. Formally, we have

$$\begin{aligned} R(t,T) &= \frac{1}{(T-t)} lnA(t,T) + \frac{1}{(T-t)} B(t,T) r(t) \\ lnA(t,T) &= ln\frac{P(0,T)}{P(0,t)} + B(t,T)F(0,t) - \frac{1}{4a^3}\sigma^2(e^{-aT} - e^{-at})^2(e^{2at} - 1) \\ B(t,T) &= \frac{1 - e^{-a(T-t)}}{a} \end{aligned}$$

The above equations are due to MathWorks (2014). One can observe that a given yield curve is always generated in completely deterministic way, only $r(t)$, the starting point, is a stochastic variable.

The stochastic discount factor for the period (t_1, t_2) could be represented as

$$D(t_1,t_2) = exp(-\int_{t_1}^{t_2} r(t)dt)$$

(Glasserman, 2004). That is, we express quantity of time t_2 at time t_1. It is a stochastic variable, since $r(t)$ is stochastic. To obtain this quantity, the obvious approach would

be to calculate the integral on the right hand side the of the expression, but that would require the stochastic rate to be simulated on a grid as fine as possible and it would include approximating another integral (Brigo and Mercurio, 2006). Instead, if we define $P(t,T)$ as a value of zero-coupon bond at time t, paying exactly one dollar at time T, it is possible to show that $P(t,T) = D(t,T)$ and $P(t,T)$ is an expectation of $D(t,T)$ under a certain probability measure (Brigo and Mercurio, 2006). The upshot of the last relationship is the possibility to calculate $P(t,T)$ analytically and we do not have to concern ourselves with on how detailed time grid to generate the short-rate path.

Chapter 4
Counterparty Credit Risk

Counterparty credit risk is faced by participants on Over-the-Counter financial markets (OTC), where there is no central authority clearing and settling the transactions. On usual exchanges such central authority exists and in case of default of a participant, the central authority steps in and settles the transaction on its behalf. On OTC no such central authority exists and in case of a default participants are more vulnerable and face losses, which can be considerably high when trading derivatives (Brigo et al, 2013). Such derivatives are often constructed to protect against market risk, but there is also significant credit exposure created. To guard against counterparty risk there is number of protective measures, such as posting collateral and marking-to-market, which is accounting technique that requires counterparties to post loss or profit made with regard to the position every day (Brigo et al, 2013). This effectively spreads the risk and mitigate the situations, where a position is settled by one large lump-sum. Counterparty risk is best defined in terms of CVA. CVA is the difference between the price of a portfolio assuming risk neutrality and portfolio with incorporated possibility of default or change in creditworthiness of our counterparties (Pykhtin and Zhu, 2007). It is the value financial markets are allocating to counterparty risk (Brigo et al, 2013). There is number of other measures of counterparty credit risk once could discuss, such as the Debit Value Adjustment (DVA), the Funding Value Adjustment (FVA).

Counterparty credit risk was not given enough attention prior to the crisis, which was also evident from the composition of Basel II. Basel II does not take care of counterparty credit risk, when it comes to capital charges, institutions were not required to set aside any capital, which would cover this risk (BCBS, 2010). This is remedied by Basel III, which has capital requirements covering counterparty risk. It approaches financial institutions from two sides, they either can have their own inter-

nal modeling system approved by a regulator ex ante or they have to use standardized calculations to quantify counterparty credit risk.

4.1 Counterparty risk and financial crisis

The underestimation of counterparty risk and its important role during financial crisis's was already pinpointed by Borio (2004). Author analyzes the dynamics of market distress and argues that the central problem of the "vanishing liquidity", which occurs during financial crisis often, is caused arguably by counterparty risk. The first phase of the financial crisis are rising asset prices over sustainable levels, risk exposures of financial institutions growing and market experiencing high liquidity. At every point in time market participants need to have an amount of cash ready to settle trades, for marking-to-market on derivatives, margin requirements. It is easier to obtain these funds in times of high liquidity. Then there comes the reversal, which leads to erratic adjustment of prices, liquidity dies-out and volatility increases. Reversal signifies that crisis is taken place. Then the erratic changes in prices, which arise when crisis erupts, impact very negatively the cash flow demand of a market participant. Marking-to-market for derivatives and margin requirements will require highly volatile cash flows to be settled. Also as market participants face losses and increasing volatility they will try to reduce their exposures, further negatively affecting prices.

However, any kind of transaction on financial markets has inherent counterparty risk as such transaction either involves exchanging cash flows for settlement and derivatives also involve large credit exposure, this requires the other counterparty to be solvent. Collateral and marking-to-market are used to limit such risk, but as there is no liquidity to find, it is going to be harder to cover marginal calls and requirements for more collateral as no other market participant wants to give up its cash, because they face the same problems. This will induce market participants to limit their trades as they want to hoard liquidity in order to protect themselves and they can not be sure if potential counterparties will be able to cover their end. Banks assume that there is higher risk of default. Interbank markets are also impacted strongly, since the lending is unsecured on this market. Result is market which will not be functioning with such depth and information asymmetry will produce uncertainty regarding creditworthiness of counterparties. Banks will not trade with each other, liquidity vanishes.

Recently Taylor and Williams (2009) find robust evidence that counterparty risk in

fact was behind widening spreads between interbank lending rates and federal funds rate in USA. The measures that US central bank introduced early, the Term Auction Facility (TAF) for instance, did not help to decrease the spread. This measure did not take counterparty risk into account.

4.2 Definition of Default

So far we have not defined what exactly constitutes a default. The default is not merely restricted to a bankruptcy, but rather when counterparty fails to honor the obligations of a contract in a meaningful way (Brigo et al, 2013). This usually includes missed payments for which remedy has not been carried out during a given period. Bankruptcy can be initiated by creditors or by counterparty itself and it is formally decided in a court of law, which decides the fate of the counterparty here on in. Bankruptcy could be understood as default in an official sense of the word, but there are as well situations when default is not necessarily caused by bankruptcy. Brigo et al (2013) gives the following six types of credit events defined by ISDA, which shall encompass the concept of default as buyers and sellers of credit default swaps understand it.

1. Bankruptcy - decided by court of law
2. Failure to pay - missed payment assumed certain restrictive assumptions hold
3. Restructuring company's debt
4. Repudiation - bankruptcy of sovereign debtor
5. Covenant violations - debt contract may include restrictions, whose breach than leads to liquidation of the debt for instance

4.3 Definition of Exposure

Alternative explanation for Counterparty risk is that if one of counterparties in a derivative contract defaults, the other counterparty faces losses in the amount equal to maintaining the same market position as before by engaging in a similar derivative contract with another counterparty (Pykhtin and Zhu, 2007). The market value of the derivative contract to the counterparty can be either negative or positive. If it is negative, then at the time of default it pays the defaulting counterparty the appropriate market value, closes out the position and at the same time engages in another

contract similar to previous, which pays our counterparty its market value. In the perfect market the loss experienced by our counterparty should be zero. If the value of contract is positive, our counterparty faces a loss equal to the market value of it, since defaulting counterparty can not honor its obligations and we have to buy similar contract for its market value in order to keep unchanged market position. The following definitions follow the rigor set up by Pykhtin and Zhu (2007).

Definition 5. Contract-level exposure at time t is defined as follows

$$E_i(t) = max\{V_i(t), 0\}$$

where $V_i(t)$ denotes the value of i-th contract at time t.

Hence contract-level exposure applies for the case of portfolio with a single derivative contract for a given counterparty. If portfolio contains more than a single contract for a given counterparty the counterparty-level exposure needs to defined.

Definition 6. The counterparty-level exposure at time t is defined as the sum of individual contract-level exposures as follows

$$E(t) = \sum E_i(t) = \sum_{i=1}^{n} max\{V_i(t), 0\}$$

where n is the number of contracts with a given counterparty.

One of the most important mitigants of counterparty risk are the netting agreements, which allow for the offsetting effect between contracts with positive and negative market value and reduce the overall exposure significantly. However, netting agreements with a given counterparty does not have to apply for all the contracts, which are held between the two counterparties. To generalize the above definition we give the definition for counterparty-level exposure under netting agreements.

Definition 7. The counterparty-level exposure at time t under netting agreements, which covers the first k of n contracts, such that $k \leq n$, is defined as

$$E(t) = max\{\sum_{i=1}^{k} V_i(t), 0\} + \sum_{i=k+1}^{n} max\{V_i(t), 0\}$$

where n is the total number of contracts.

For full netting agreements the previous definition also applies if we set $k = n$.

Counterparty risk is mostly concerned with how these exposures change with respect to time t, t being somewhere in the future, by means of exposure profiles. Exposure profiles are measures, which individually show different aspects of the exposure distribution at a given set of simulation dates, sometimes with respect to different probability measures. Most important profiles are The Expected Positive Exposure (EPE), The Credit Value Adjustment (CVA), The Effective Expected Positive Exposure (EEPE). These are important for regulatory purposes. There are also other such as Peak Expected Exposure (PEE), which is the harshest one as it serves as a bound on all the others. PEE is defined as a maximum value of the expected distribution, i.e. its peak. This measure is then constant in time and for all t it is equal to the peak. Following expansion of EPE's and EEPE's definitions from Ghamami and Zhang (2013) offers itself readily.

Definition 8. The expected positive exposure at a counterparty-level under netting agreements is defined as follows

$$EPE = \int_0^T E[CE(t)]dt$$

where $E[CE(t)]$ is the expected value of counterparty-level exposure under netting agreements at time $t \geq 0$ and $T > 0$ is equal to the maturity of contract in portfolio with the furthermost date of expiration.

Next we define EEPE.

Definition 9. The EEPE at counterparty-level under netting agreements is defined in continuous and discrete versions followingly

$$EEPE_d = \sum_{i=1}^{n} max_{1 \leq j \leq i} CE(j) \triangle_i$$

$$EEPE_c = \int_0^T max_{0 \leq s \leq t} CE(s)dt$$

where $\triangle_i = t_i - t_{i-1}, i = 1, ..., n$ and $0 = t_1 < ... < t_n = T$.

EEPE is more strict measure than EPE, since it accounts for the roll-over risk (Ghamami and Zhang, 2013). Such risk results from the assumption that when more

short-term securities in our portfolio mature before time T, bank would in reality replace them with the new contracts and hence bank's exposure would not be as much affected in a downward manner due to close-out of these positions. EPE does not take into account the replacement of maturing contracts in the portfolio.

4.4 PDS vs DJS Sampling

When studying the simulation of Brownian motion, the initial recurrence given for this task was of sequential nature. Assuming a time grid $0 = t_1 < ... < t_n = T$, value of $W(t_i), i > 0$ one samples conditionally on the value of $W(t_{i-1})$. Such an approach to simulation is called the Path-Dependent Simulation (PDS). Under PDS, the values of Brownian motion at different points in time are always correlated to previous values in the simulation. On the other hand, generating all the values of $W(t_i), i = 1, ..., n$ given only initial value $W(t_0)$ yields very different results. This time, values of Brownian motion at different points in time will be uncorrelated. Such an approach to simulation is called the Direct-Jump to Simulation date (DJS). We gave a specific recurrence for this case as well.

Both terms, DJS and PDS, were first coined by Pykhtin and Zhu (2007), and the authors are first ones, who point out the its importance with relation to credit exposure simulation. One should as well note that both approaches yield appropriately distributed Brownian motion. We will later find out that these approaches yield very different results, when it comes to variance of the estimator.

4.4.1 Simulation of EPE and PDS vs DJS

More practically, Ghamami and Zhang (2013) show importance of the distinction between these two approaches by considering the simulation of EPE. Again assume a time grid $0 = t_1 < ... < t_n = T$, Brownian Motion $\{W(t_i), i = 0, .., n\}$ and contract-level exposure $V(t_i)$ driven by credit exposure process $\{C(t_i), i = 0, ..., n\}$ as such $V(t_i) = max\{C(t_i), 0\}$. The exposures are some function of the Brownian motion, so that $V(t_i) = f(W(t_i))$.

The integral defined in .... can be approximated by right Riemann sum approximation as follows

$$EPE = \sum_{i=1}^{n} E[V(t_i)] \triangle_i$$

$$\widehat{EPE} = \sum_{i=1}^{n} \bar{V(t_i)} \triangle_i$$

where $\triangle_i = t_i - t_{i-1}$, $\bar{V(t_i)}, i = 1,...,n$ is the sample mean of m simulated values of the exposure at a given date t_i. Next to obtain a variance of this estimator, which is a measure of its accuracy, Ghamami and Zhang (2013) show that

$$Var(\widehat{EPE}) = Var(\sum_{i=1}^{n} \bar{V(t_i)} \triangle_i) = \sum_{i=1}^{n} \frac{Var(V(t_i))\Delta_i}{m} + \frac{2}{m} \sum_{i<j} \sum Cov(V(t_i), V(t_j))\Delta_i \Delta_j$$

.

Under DJS approach $Cov(V(t_i), V(t_j)) = 0, i, j = 1,...,n$, while under PDS approach $Cov(V(t_i), V(t_j)) \neq 0, i, j = 1,...,n$. It is straightforward to see why DJS approach yields uncorrelated exposures at a given two distinct days, these exposure are in no way related to each other, since their underlying risk factors, $\{W(t_i), W(t_j)\}$ are uncorrelated. For the PDS approach, Ghamami and Zhang (2013) prove that under few mild assumptions, the covariance of exposures is positive, that is for any $t_i, t_j, i, j = 1,...,n$, we have $Cov(V(t_i), V(t_j)) > 0$.

The assumptions differ based on nature of the credit exposure, since in some transactions, unilateral, the counterparty risk is faced only by one counterparty in contract, and in other transactions the risk is bilateral, such as interest rate swaps. The assumptions include the monotonicity of payoff function in bilateral transactions, the numeraire and the deterministic short rate in unilateral transactions.

Hence, due to zero covariance terms, the DJS approach seems to be far superior to PDS in terms of accuracy, that is variance. Both approaches, DJS and PDS, in this case yield identical computational demands as the same number of values of Brownian motion is generated in both cases. Hence, DJS seems to be far superior to PDS.

4.4.2 Path-dependent Processes

The discussion so far only considered path-independent processes in the form of credit exposure of path-independent derivative. For path-dependent derivatives the DJS approach has to be modified and its superiority seen thus far may not hold. The exposure of path-independent derivative at any given time may depend only on the value of risk factor, Brownian motion, at that time. In contrast, the exposure of path-dependent at time t_i depends on the whole realized path of Brownian motion up to

and including time t_i, that is $V_i = f(W(t_0), ..., W(t_i))$ for path-dependent derivative and $V_i = f(W(t_i))$ for the latter one.

No need to modify the PDS approach, but DJS again requires $Cov(V(t_i), V(t_j)) = 0, i, j = 1, ..., n$. To obtain $\{W(t_i), i = 1, ..., n\}$ we generate sequentially the full path between $W(t_0)$ and $W(t_i)$ and save only the last value. DJS is just a combination of n runs of PDS algorithm for different t_i. Under DJS we need to generate exactly $n(n+1)/2$ values of Brownian motion, while PDS again requires only n values.

The complexity of variance under both approaches has not changed. Consequently, both approaches seem to offset each other up to a certain degree with its advantages and disadvantages. For the DJS the advantage in lower variance is offset by the need to generate more random numbers to achieve zero covariance terms.

4.5 DJS vs stochastic discounting

In case of EPE, the requirement for DJS approach was to make the covariance terms of exposures equal to zero. However, CVA requires calculation of the discounted exposure given that the underlying asset is modeled within a stochastic interest rate model, which interest rate swap is. The discount factor, $D(0, t_2)$, which discount a quantity from time t_2 to time 0, is path-dependent. This discount factor is given as $exp(-\int_{t_1}^{t_2} r(t)dt)$, where $r(t)$ is the short-rate modeled by a given short-rate model, in our case it is the Hull-White model, under the risk-neutral measure (Brigo and Mercurio, 2006).

If again the time grid assumed is given as $0 = t_1 < ... < t_n = T$, then for any $t_i, i = 1, ..., n$ the discount rate evidently depends on the realized path of the short-rate from 0 to t_i. Now, assume a portfolio of interest rate swaps, whose exposure depends on the risk factor S, $\{S(t_i), i = 1, ..., n\}$. These risk factors in turn depend on the short rate $r(.)$, $\{r(t_i), i = 1, ..., n\}$, i.e. $S(t_i) = S(r(t_i))$. The risk factor for an interest rate swap would be a yield curve.

Generating a discounted exposure at time t_i consists of generating the short-rate path $(r(t_0), ..., r(t_i))$ and the risk factor at t_i only, $S(t_i)$. Then we calculate the required discounted exposure given as $D(t_i)V(t_i)$. At time $t_j, j > i$ sampling from $D(t_j)V(t_j)$ is possible by generating the short-rate from $r(t_i)$ and continuing up to $r(t_j)$ (Pykhtin and Zhu, 2007). Then again calculating the risk factor and discounted exposure eventually. However, this time we have that $Cov(D(t_i)V(t_i), D(t_j)V(t_j)) \neq 0$, since discount factors for both exposures and in part even exposures depend on the common

set of short rates.

DJS approach is valid then if for any discounted exposure $D(t_k)V(t_k), t_k, k = 1, ..., n$ the simulation is carried out by always generating the new short-rate path from the beginning. Hence, DJS under stochastic discounting is similar to a DJS for path-dependent derivative.

Valid assessment of the computational intensity is considering again the trade-off between variance and the computing time. Ghamami and Zhang (2013) suggest

$$variance\, x\, E(computing\, time)$$

This is due to the discussed expectation that DJS is to have higher computational demands, but lower variance than what PDS approach offers.

4.6 CVA

CVA is positive subtraction from value of risk-free portfolio, which accounts for portfolio's counterparty risk. This proposition is formally proved by Brigo et al (2013). He proves that the following applies

$$\widehat{NPV}(t,T) = NPV(t,T) - CVA$$

Where the first term on the right-hand side is a discounted value of risk-free portfolio and the left-hand side is equal to the value of risk-free portfolio. The second term on the right-hand side is then the Credit Value Adjustment.

4.6.1 Definition of CVA

There is a number of ways how to define this quantity CVA. We can define CVA also as a running spread (Vrins, 2011). We will proceed from the most general exposition to an analytical formula, which will be then utilized during practical simulations. First of the assumptions to be fulfilled is the assumption that, we, as a counterparty are default-free and all the other counterparties, we are in business with, are not. We are always calculating the counterparty risk from the point of view of our default-free counterparty. This assumption may not be completely realistic, but is crucial for reducing the complexity of the task ahead. It is then possible to define Unilateral

CVA. On the other hand, bilateral CVA would take into account also the probability of our default. The following definition is due to Pykhtin and Zhu (2004).

Definition 10. Unilateral CVA is defined as a risk-neutral expectation of discounted loss due to a default of given counterparty. The discounted loss is given as follows

$$L = 1_{\{\tau \leq T\}}(1-R)\frac{B_0}{B_t}E(\tau)$$

and unilateral CVA is given as follows

$$CVA = E[L] = (1-R)\int_0^T E[\frac{B_0}{B_t}E(\tau)|\tau = T]dF(t)$$

T is the time horizon equal to a maturity of an instrument with the longest maturity in the portfolio, B_t is the value of a bond contract of face value 1 at time t, $E(t)$ is the exposure at time t. $1_{\{\tau \leq T\}}$ is the event indicator, $F(t)$ is the risk-neutral probability of counter-party's default at time t, τ denotes the supposed default time of the counterparty. The expectations is taken with respect to the risk-neutral density and finally R is the recovery rate.

The above quantity is always positive (Brigo et al, 2013). It could be argued that the value of financial instrument with default-free counterparty is always higher than value of financial instrument with default-able counterparty. Also we can see that if the default time exceeds the given time horizon, there will be no less experienced by us.

Next proposition is from Brigo et al (2013) and it redefines CVA to a simpler form, which will be than picked apart with another assumptions.

Proposition 11. *The Credit Value Adjustment (CVA) discounted to present can be given as follows*

$$CVA = E[1_{\{1<\tau\leq T\}}(1-R)D(\tau)E(\tau))]$$

The expectation is taken with respect to the risk-neutral measure, R is the recovery rate, τ is random variable denoting the default time of counterparty, $D(...)$ is the factor discounting the loss to present and $1\{.\}$ is the event indicator.

The derivation of the previous formula is described in detail in Brigo et al (2013).

4.6.1.1 CVA on a discrete time grid

Since we can not practice in the continuous time in the real world, we have to derive a similar formula applicable on a discrete time grid. Assume one such discrete time grid $0 = t_0 < t_1 < ... < t_n = T$. Brigo et al (2013) discretizes and approximates the previous expression as follows

$$CVA = (1-R)\sum_{i=1}^{n} E[1_{\{t_{i-1}<\tau\leq t_i\}} D(\tau)E(\tau)]$$

This gives a way how to practically estimate CVA for every given time bucket, of course assuming the default occurs inside that bucket. It would be more useful to assume the default does not occur somewhere inside the time bucket, but rather at the end of the particular time bucket. If the time grid is fine enough, this assumption is convenient and does not lead to serious over-simplification (Brigo et al, 2013). Consequently, we can re-write the expression in previous proposition to yield

$$CVA = (1-R)\sum_{i=1}^{n} E[1_{\{t_{i-1}<\tau\leq t_i\}} D(t_i)E(t_i)]$$

This representation is still not sufficiently tractable for us. We need a representation, which will allow us to model the probability of default independently of the modeling of the exposure process. The previous representation would require sophisticate default model in order to model the value of the instrument together with the default. Hence, next step is to assume the independence of size of the exposure and the default time τ. This assumption is justified for the case of interest rate swaps as we will find out later. We can consequently follow Brigo et al (2013) and factor out the probability of default and write

$$CVA = (1-R)\sum_{i=1}^{n} F(t_i)E[D(t_i)E(t_i)]$$

By assuming independence between exposure and the default variable τ, the wrong-way risk is disregarded. It is possible for exposure to increase with increase in the probability of default of the counterparty. That is called the wrong-way risk. The risk would be right-way if exposure instead tends to decrease under otherwise same conditions. However, wrong-way risk is assumed to be not as important for interest rate derivatives Pykhtin and Zhu (2007). Next we will describe two approaches how to make simulation of CVA potentially more efficient.

4.6.2 Efficient MC Simulation of CVA

It is noticeable that EPE and CVA are both weighted sums of expected exposures. Using the property of iterative expectations and Ghamami and Zhang (2013) and the results from last section we obtain

$$CVA = (1-R)E[E[1\{\tau \leq T\}(D(\tau)V(\tau))|\tau]] = \int_0^T E[V(t)D(t)]dF(t)$$

where F is the cumulative distribution function of the default variable τ. For CVA the weights are probabilities of default for a given sub interval, while for EPE weighs the expected exposures using the length of particular sub interval. The right Riemann sum approximation for the previous integral is

$$\widehat{CVA} = \sum_{i=1}^{n} \overline{D(t_i)V(t_i)}\Delta F(t_i)$$

where $\Delta F(t_i) = F(t_i) - F(t_{i-1})$, $\overline{D(t_i)V(t_i)} = \sum_{j=1}^{m} \frac{D(t_{ij})V(t_{ij})}{m}$ is sample mean of m discounted exposures. Hence we have n time buckets and m scenarios. We draw a i.i.d. sample of size m at every time bucket. This estimator likely produces biased results due to time-discretization (Glasserman, 2004). The exact joint distribution of stochastic process is sometimes unachievable and approximate solutions, such as Euler scheme, do away with bias only as the time step converges to zero.

Effectively, one can choose between minimizing variance and minimizing bias. If the number of scenarios m increases, the variance decreases as can be seen from the variance of estimator for EPE. Then under fixed computational budget the number of time buckets, n, has to be reduced, which in turn increases bias (Glasserman, 2004). Time-discretization bias is somewhere around $\frac{1}{n^\beta}, \beta \geq 1$ (Glasserman, 2004). One performance measure that takes into account bias and variance at once is Mean Square Error (MSE), previous literature points to importance of choosing the right number of scenarios and time buckets based on minimizing the MSE.

4.6.2.1 Stratified CVA estimator

Using the discussed stratification, one can obtain the unbiased estimator as follows

$$CVA \quad = \quad (1-R)E[1\{\tau \leq T\}(D(\tau)V(\tau))] = (1-R)\sum_{i=1}^{K} E[D(\tau)V(\tau)|\tau \varepsilon A_i]$$

$$= \sum_{i=1}^{n} \overline{D(\tau_i)V(\tau_i)}\Delta F(t_i)$$

where $A_i, i = 1,..,n$ is appropriately defined a set of disjoint strata and τ is the stratification variable. Suppose we allocate the number of scenarios for each strata as $m_i, i = 1,..,n$ so that following applies $\sum_{i=1}^{K} m_i = M$, where M is the total number of scenarios. K is the number of strata and effectively the number of time buckets. The average inside the previous expression is obtained by drawing a conditional sample from $F(\tau|\tau \in A_i)$ of size m_i and for every point in this sample we draw $D(\tau_{ij})V(\tau_{ij})$, then form a sample mean as described. This estimator is due to Ghamami and Zhang (2013).

As opposed to previous estimator, here the sampling of discounted exposures is done inside the individual time buckets randomly. Also in comparison, here additional computational effort is required for generating conditional samples from distribution of default probabilities. The size of samples from strata can be selected arbitrarily and proportional sampling $m_i = MF(\tau\varepsilon A_i)$ was seen to only decrease variance with respect to non-stratified estimator. The optimal allocation in the form of $m_i = \sigma_i F(\tau\varepsilon \in A)$, where σ_i is the variance within i-th stratum is not possible directly as the true value of this variance is unknown.

4.6.2.2 Implementation complications

To utilize stratified sampling for CVA simulation, one would have to able to draw conditional samples from the default probability distribution $F(t)$. This would require in turn more sophisticated simulation of probability of default so as to obtain analytical representation of the default probability distribution. We will use bootstrapping techniques to obtain default probabilities from credit default swaps of given counterparties. This is more elementary approach and we will not be able to obtain default probability distribution from which conditional samples can be drawn. Also available data are not as detailed and hence do not contain as much information in order to fully utilize the advantages of stratification variable.

Another complication is that at our time horizon T up to which simulations are carried out, the default probability distribution does not likely reach the value of 1. That complicates the setting number of observations allocated to each stratum.

On the other hand, Ghamami and Zhang (2013) do not mention the superior computational demands of sampling the exposures and that this stratified estimator may

not lead to much reduction in variance. In our framework with analytical representation of CVA the importance of modeling of default is not as high. Generating exposures constitutes majority of the computational intensity.

4.6.3 MSE Minimization

The previous literature in Ghamami and Zhang (2013) employs a Mean Square Error (MSE) minimization in order to determine the optimal selection of number of paths, m, and the number of valuations points $n$in order to reduce variance and bias.

We have seen that it is not generally feasible to generate stochastic processes with the exact distribution of its theoretical counterparts. Rather we have to get accustomed to using methods such as the Euler approximation, which under fixed n does not reproduce the exact joint distribution one would require. And so under fixed n our estimator is biased, due to the so-called time-discretization bias, which is of order $\frac{1}{n}$ under the right Riemann sum approximation (Glasserman, 2004). However, we are guaranteed that the estimate would approach its true value as $n \to \infty$.

On the other hand, under fixed budget selecting higher number of valuation points reduces the number of replications or paths, m, we can generate.The number of replications determines the size of the variance (Glasserman, 2004). Hence, one can see the trade-off between selecting the number of valuations points in order to reduce bias and number of replications to reduce variance. MSE minimization should in theory yield optimal combination of both numbers. The MSE is defined as follows.

$$MSE(\hat{\theta}) = Var(\hat{\theta}) + (E(\hat{\theta}) - \hat{\theta})^2$$

where $\hat{\theta}$ is the estimator of the desired quantity.

The spacing and number of valuation points is in risk management generally determined by the prevailing practice and is not clearly justified by particular reasons. The majority of valuations dates are closer to present, and then their density decrease with increasing time. This is probably due to the observations that discount factor reduces the importance of an observation with respect to increasing time.

4.6.3.1 MSE minimization for EPE

To derive a minimization problem involving number of scenarios m and number of valuation points n, we need to represent MSE presented above in terms of these two

unknowns. We cannot formulate this problem in terms of CVA estimation, but rather in terms of EPE. The reason is that the form of CVA estimator does not allow for such analytical manipulation as the form of EPE estimator does. We already specified of the variance of PDS estimator of EPE as follows

$$Var(\hat{\theta}_{PDS}) = \sum_{i=1}^{n} \frac{Var(V(t_i))\Delta_i}{m} + \frac{2}{m} \sum_{i<j} \sum Cov(V(t_i), V(t_j))\Delta_i \Delta_j$$

The variance of DJS estimator of the same quantity is on the other hand

$$Var(\hat{\theta}_{DJS}) = \sum_{i=1}^{n} \frac{Var(V(t_i))\Delta_i}{m}$$

First, Ghamami and Zhang (2013) show that the variance of DJS estimator $\hat{\theta}_{DJS}$ can be given an upper bound of $\frac{cT^2}{mn}, c \in R$. T is defined as the time horizon and we have to assume the exposure process to have finite second moment and time grid with equally spaced valuation points. Using the Landau symbol we see that the variance is of order $O(\frac{1}{mn})$. For bias, we already specified it is of order $O(\frac{1}{n})$. But for MSE we have to include the squared bias. Hence, we need to solve the following optimization problem

$$min_{m,n}(\frac{1}{mn} + \frac{1}{n^2}), s = mn$$

where s is our computational budget (Ghamami and Zhang, 2013).

Possible solution of the above problem is to set $m = 1, s = n$. This shall be interpreted as devoting all of the budget to a valuation points and set number of scenarios to once.

For PDS estimator, $\hat{\theta}_{PDS}$, the only difference lies in the complexity of variance, where covariance terms of exposures are dependent and the variance is of order $O(\frac{1}{mn} + \frac{1}{m})$. The last term accounts for the covariances. Ghamami and Zhang (2013) show that similar optimization problem for the PDS estimator yields not so clear results. Optimal selection is estimated to be as $n \cong s^{\frac{1}{3}}$ and $m \cong s^{\frac{2}{3}}$.

Obviously, these optimization problems are solved only approximately. However, the results for DJS approach are very unambiguous.

4.6.3.2 MSE minimization vs Exposure Sampling

The MSE minimization is arguably not as useful as one would hope if exposure sampling is taken into account. We found out that under DJS approach for path-

independent derivatives all of the fixed budget is to be devoted to a valuation points, n, while under PDS approach weighted division between the two is advised.

The previous results apply only for estimation of EPE. Ghamami and Zhang (2013) also study the optimal selection of m and n for estimation of EEPE. The optimal selection of m and n in the latter case is of course different. This results from the form of EEPE, where the mathematical operator $max\{...\}$ introduces new source of bias. Ghamami and Zhang (2013) show that in order to minimize MSE of EEPE estimator, one selects m and n according to following optimization problem

$$min\,(\frac{a}{mn}+\frac{b}{m}+(\frac{c}{m}+\frac{d}{n})^2); s = mn; a,b,c,d \varepsilon R$$

What we arrive at in both cases is the different optimal number of valuation points n and replications m. And we are dealing with only two different measures of counterparty risk. In reality there are various other measures one desires to estimate.

In our case, the most computationally intensive part is generating the portfolio's exposure. The minimization approach discussed here implicitly assumes that one generates the exposures specially for every given measure, and most likely at different points in time. This is not an effective approach, the exposure sampling offers a different view. Rather than selecting n arbitrarily for every different measure, under exposure sampling we would first generate the distribution of exposures. Consequently, to estimate any measure we would draw a sample from this distribution. This is done easily as every measure in some form depends on expected positive exposure of the portfolio. Most importantly, exposure sampling is also subject to a time-discretization as any other simulated stochastic process in practice. Hence it will be possible to draw samples from this distribution only at finite set of points, without the possibility to arbitrarily pick a valuation date. We will have to generate the exposure of a portfolio only once and yes the specific calculation of EPE, EEPE and CVA will not be as effective, but it will be more than offset with less demanding exposure sampling. MSE minimization in other words is only a micro-optimization.

4.6.3.3 Practical simulation of CVA

For practical simulations and calculation of CVA we will first employ naive approach, when it comes to m and n selection and then try to select the constants given the results from the MSE minimization in EPE's case. As explained, the form of CVA's variance is such that it is not possible to derive an upper bound as for EPE. However, the independence of exposures assumption is still satisfied for the DJS approach under

stochastic discounting and the results for DJS approach are so clear-cut that it is likely it could work even for the case of different probability measure.

Hence, under DJS we will also try to spend the whole computational budget on making the valuation dates as close as possible. It is important to note that making the grid as fine as possible is specifically due to the bias part of the MSE as we have discussed. Since in reality we never know the true value of an estimate, we can not estimate the amount of bias and only calculate the reduction in variance. That does not have to be as large with all of the budget spent on n.

Chapter 5
Practical Simulations

5.1 Swap portfolio description

We have a portfolio of swaps with 5 different counterparties and for each of them the counterparty risk measures are calculated. The data are obtained thanks to MathWorks (2014). The statistics of the portfolio are presented in the following table. Counterparty paying the fixed rate is denoted as 1. The decimal number in columns Leg Rate Receiving/Leg Rate Paying denotes the floating rate that counterparty pays or receives. The integer in the same columns denotes the number of basis points over the zero-rate, which counterparty pays or receives through the fixed payment stream.

The assumption of full netting agreements is taken, since the counterparty risk of swap portfolio without netting agreements is just a sum of counterparty risk of individual swaps (Brigo et al, 2013). Such a case would be better studied with a single swap. The individual swaps are all 1-year swaps, that is the floating rate is reset every 12 months. However, they all have different maturities and reset dates are not homogenous, but are spread over the whole year. We will have to employ a number of approximations in order to find appropriate floating rates for differing reset dates.

5.2 The default probability distribution estimation

The default probabilities, which are used in estimating the CVA, are stripped from credit default swaps (CDS) for a given counterparty. This is the most convenient way to obtain default probabilities under a risk-neutral measure. The data on CDSs are

CounterpartyID	Principal	Maturity	Fixed Payer/Receiver	Leg Rate Receiving	Leg Rate Paying
1	629468	18-Feb-2012	1	0,0387	0
1	969469	23-Sep-2009	0	78	0,0400
1	880538	03-Dec-2009	1	0,0388	39
1	512488	23-Jul-2012	0	12	0,0402
1	873121	15-Jun-2009	0	56	0,0386
1	675866	09-Aug-2011	1	0,0399	22
2	660412	12-May-2012	0	8	0,0395
2	815297	04-Dec-2014	1	0,0429	11
3	353968	07-Oct-2011	1	0,0414	36
3	432644	13-Feb-2012	0	24	0,0403
3	397446	12-Jul-2010	1	0,0428	78
4	918177	20-Jul-2014	1	0,0475	74
4	712034	30-Jan-2013	1	0,0436	49
4	513745	27-Aug-2011	1	0,0439	77
4	937895	12-Nov-2009	0	36	0,0373
4	464379	26-Nov-2013	1	0,0414	7
4	817900	06-Mar-2012	1	0,0404	22
4	535334	02-Jun-2009	1	0,0368	17
5	813450	28-May-2009	1	0,0361	10
5	441321	10-Apr-2009	0	87	0,0393
5	774308	15-Aug-2013	0	70	0,0463
5	361971	08-Jan-2009	1	0,0363	23
5	443131	21-Dec-2010	0	72	0,0423
5	440712	17-Sep-2013	1	0,0473	82
5	860714	26-Oct-2010	0	16	0,0384
5	946948	11-Dec-2009	0	13	0,0367
5	438313	14-Nov-2012	1	0,0443	52
5	604967	08-Jun-2013	1	0,0411	13
5	688948	28-Apr-2010	1	0,0390	32
5	662293	05-Jun-2014	1	0,0448	46

Table 5.1: Description of the portfolio

yearly and include 5 observation up to the settle date for every counterparty in the portfolio. It is obvious that such few observations do not contains as much information as one would like about the default distribution of given counterparty.

5.3 Practical simulation steps

When carrying out our practical simulations. we follow closely MathWorks (2014). The first step is to select an initial zero-rate curve. This is an arbitrary choice. To select valuation points we proceed with 12 dates 1 month apart for the first year and after a first year the frequency drops to a one date every 3 months for another 6 years.

The valuation points are not in line with the individual reset dates, hence an approximation is to be carried out. The first part is to calculate for each valuation date the previous reset date, this needs to be done to appropriately value the floating part of the swap. As noted the reset dates are not homogenous and differ from valuation dates. Hence the 1-year floating rate for each valuation date is approximated by interpolating between the two nearest interest rate curves, which squeeze the last reset date corresponding to the valuation date between themselves. Then the valuation of individual swaps is carried out with respect to the interest rate curve at a given valuation date.

Next task is to calibrate the Hull-White model we are using, this is done by specifying constants α and σ. We select the α to be 0.2 and $\sigma$to be 0.015. Then for each valuation date the interest rate curve is generated and this is repeated according to the number of scenarios specified. The number of scenarios for PDS approach will vary between 1000 and 16 thousand.

The number of scenarios for PDS approach will be 1 000, 4 000, 16 000. We always increase the number of scenarios by multiplying the previous number by 4. As was pointed out earlier, Monte Carlo method provides the convergence rate of order $\frac{1}{\sqrt{n}}$. This means, if we increase the number of scenarios by a factor of four, we should be able to obtain estimates with standard errors cut in half. However, with higher number of scenarios generated, we will as well see the computing time to increase.

To estimate the variance we iterate 100 times the whole computation with a given number of scenarios. It would be desirable to iterate at least 1 000, but due to the limited computing power we have to scale back. The DJS approach and also the PDS approach with 16 000 scenarios would take too long to finish. However, according to

our observations the estimates of variance are quite stable and no significant change is observed, when iterating 1 000 times as opposed to 100 times.

5.4 Results

The results are summarized in Tables 5.2 and 5.3. The table 5.2 contains statistics on computing time, variance and the compound statistic. For variance and computing time, we divide the values received due to different approaches by the value from PDS approach with 1 000 scenarios. This allows to see by how many orders the values of various approaches differ.

We can see that DJS approach yields estimates with variance lower by many orders of magnitude than PDS approach with the same number of scenarios. However, the reduction in variance varies significantly between the counterparties. For two counterparties the variance drops by 11 orders of magnitude and for one the variance drops by 19 orders of magnitude. Variances are often higher than the estimate alone, which implies we can not be sure, where the true value lies. The likely cause is the insufficient number of repetitions. The reduction in variance is especially high for counterparties with high variances of their CVA estimates under PDS. On the other hand, the estimates for counterparty, which is already relatively precise under PDS, does not improve as much under DJS.

Although, DJS produces estimates with lower variance, we also observe significant increase in computing time as opposed to PDS with the same number of scenarios. In the next part, we utilize the discussed metric accounting for both the computing time and for variance. The increase in computing time under DJS is by 13 orders of magnitude. One can observe that gains in preciseness are fully wiped out by the increased computing time in four cases out of all five. It is important to note that the results on computing time in our case serve as an upper bound. It would be possible to build the computational system better in order to fit more the DJS approach (Albanese et al, 2011).

We also tested the PDS approach with 4 000 and 16 000 scenarios. The errors in estimates improved more than expected. When going from 1 000 to 4 000 scenarios and from 4 000 scenarios to 16 000 scenarios, we saw standard errors to be cut by 3-4 orders of magnitude, respectively by 15 orders of magnitude. The computing time increases by 3 orders of magnitude, respectively by 11 orders of magnitude. One could observe that convergence rate of Monte Carlo method significantly underrates

Table 5.2: The compounded results

DJS, 1000 scenarios	Counterparty 1	Counterparty 2	Counterparty 3	Counterparty 4	Counterparty 5
Mean estimate	507,54	1026,63	360,37	3184,91	1491,58
Variance	132,43	252,50	36,10	2507,61	711,00
Expected computing time	361,92	361,92	361,92	361,92	361,92
PDS, 1000 scenarios					
Mean estimate	507,61	1178,91	279,25	5051,51	1958,05
Variance	357,55	2832,33	29,24	49332,36	8070,95
Expected computing time	26,44	26,44	26,44	26,44	26,44
PDS, 4000 scenarios					
Mean estimate	509,07	1179,85	279,60	5067,58	1961,39
Variance	78,64	803,17	7,89	13849,46	2344,46
Expected computing time	79,15	79,15	79,15	79,15	79,15
PDS, 16000 scenarios					
Mean estimate	507,54	1177,25	279,81	5051,10	1956,07
Variance	18,57	189,88	1,64	3206,25	534,21
Expected computing time	303,64	303,64	303,64	303,64	303,64

Variance x E(computing time)	Counterparty 1	Counterparty 2	Counterparty 3	Counterparty 4	Counterparty 5
DJS, 1000 scenarios	47927,04	91384,42	13065,57	907551,17	257325,72
PDS, 1000 scenarios	9453,27	74884,90	773,19	1304314,67	213390,53
PDS, 4000 scenarios	6224,90	63574,00	624,88	1096242,23	185574,01
PDS, 16000 scenarios	5638,01	57654,56	498,39	973557,63	162208,29
Variance	Counterparty 1	Counterparty 2	Counterparty 3	Counterparty 4	Counterparty 5
DJS, 1000 scenarios	2,70	11,22	0,81	19,67	11,35
PDS, 1000 scenarios	1	1	1	1	1
PDS, 4000 scenarios	4,55	3,53	3,70	3,56	3,44
PDS, 16000 scenarios	19,26	14,92	17,82	15,39	15,11
Computing time	Counterparty 1	Counterparty 2	Counterparty 3	Counterparty 4	Counterparty 5
DJS, 1000 scenarios	13,69	13,69	13,69	13,69	13,69
PDS, 1000 scenarios	1	1	1	1	1
PDS, 4000 scenarios	2,97	30,38	0,30	523,82	88,67
PDS, 16000 scenarios	11,48	11,48	11,48	11,48	11,48

Table 5.3: Table with scaled variance, computing time and the compounded measure of computational intensity

the actual reduction in variance realized, when we increase the number of scenarios generated. We would expect the reduction in variance by 2 orders.

The unbiasedness of estimates is impossible to determine. At least we observe that estimates from both approaches are located always with the bound, which could be deduced from any of the variances. However, for two counterparties, the DJS approach produces markedly different estimates. They are always within the bounds given a variance of the PDS estimator. It is not clear, if it is bias or insufficient number of scenarios.

Now we consider the discussed measure

$$variance * E(computing time)$$

We look if the gains in preciseness are really overshadowed by the increase in computing time. The expected computing time is calculated as an average of computing times for each iteration.

The only counterparty for which DJS approach yields comparatively the best results is the counterparty number 4. In this case the absolute value of the product of variance and computing time is the lowest by about 20%. In all other cases, PDS approach with 16 000 scenarios is the most advantageous from the computational standpoint.

Chapter 6
Conclusion

Counterparty credit risk has seen its relevance increase after the last financial crisis. It had been the reason for majority of losses occurring during this period (BIS, 2011). These losses were caused by changes in creditworthiness of counterparties, not their actual defaults. One could ascribe such high losses to inability a properly simulate the underlying risk. Simulation of counterparty credit risk is very computationally intensive as we discussed. Hence, it is important to study the effectiveness of methods, which can decrease such intensity.

We have compared two fundamentally different approaches for stochastic modelling, which surfaced with connection to counterparty risk in Pykhtin and Zhu (2007). These two approaches, PDS and DJS, were already extensively studied by Ghamami and Zhang (2013) for the case of individual path-independent derivatives. In our case the test subject is the portfolio of swaps. We calculate portfolio's CVA and the effectiveness of both methods. Although a swap is not a path-dependent derivative, it effectively becomes one when interest rate is simulated within a stochastic model, such as the Vasicek model. The short-rate is path-dependent inside this model. DJS is proved to be significantly more efficient, when it comes to path-independent processes (Ghamami and Zhang, 2013). However, for path-dependent processes other factors come into play and DJS approach may not have to be as effective compared to PDS approach. To our best knowledge, there is no published work that deals with the problems of effective simulation of counterparty credit risk for portfolio of swaps or any kind of portfolio in fact.

We find that as expected the DJS approach yields more precise results than PDS approach for the same number of scenarios. However, the results are not uniform and for some counterparties the preciseness is more pronounced. One counterparty sees its variance to be lower by 19 orders of magnitude, another by 11 orders of magnitude

under the DJS approach.

Also as expected, the computational intensity of DJS approach is higher than that of PDS approach. To evaluate both approaches accordingly, we used a criterion, which accounts for both preciseness and also expected computing time. Only in one case, that is for a given single counterparty in the portfolio, we see DJS approach to consistently outperform the PDS approach. For any other counterparty in the portfolio, the higher computing time under DJS approach does not justify the higher preciseness of DJS approach. PDS outperforms DJS in majority of cases. Interesting finding is that PDS approach yields more precise estimates than would be predicted by the convergence rate of Monte Carlo method given increasing number of scenarios.

We have not been able to apply any variance-reduction methods using the analytical approximation from Brigo et al (2013). Also as argues Brigo and Mercurio (2006), it is always useful to include a no-arbitrage argument as a control variable. Applying both of these approaches would require more sophisticated stochastic model that which we used in this thesis. This could be a useful direction for a future research.

Bibliography

Claudio Albanese, Toufik Bellaj, Guillaume Gimonet, and Giacomo Pietronero. Coherent global market simulations and securitization measures for counterparty credit risk. *Quantitative Finance*, 11(1):1–20, 2011. URL `http://EconPapers.repec.org/RePEc:taf:quantf:v:11:y:2011:i:1:p:1-20`.

A. Antonov, S. Issakov, and S. Mechkov. Algorithmic exposure and cva for exotic derivatives. Available at SSRN: http://ssrn.com/abstract=1960773 or http://dx.doi.org/10.2139/ssrn.1960773, November 17 2011.

Basel Committee on Banking Supervision. Basel iii: A global regulatory framework for more resilient banks and banking systems, December 2010.

Claudio E. V. Borio. Market distress and vanishing liquidity: anatomy and policy options. BIS Working Papers 158, Bank for International Settlements, July 2004. URL `http://ideas.repec.org/p/bis/biswps/158.html`.

Damiano Brigo and A. Alfonsi. Credit default swaps and option pricing with ssrd stochastic intensity and interest-rate model. *, Finance & Stochastics*, Vol. IX(1), 2005, 2005.

Damiano Brigo and Fabio Mercurio. *Interest Rate Models: Theory and Practice - with Smile, Inflation and Credit*. Heidelberg, 2nd edition, 2006.

Damiano Brigo, Massimo Morini, and Andrea Pallavicini. *Counterparty Credit Risk, Collateral and Funding: with Pricing Cases for all Asset Classes*. Wiley, 2013.

McKinsey Company. Getting to grips with counterparty risk. *McKinsey Working Papers on Risk*, 20, June 2010.

Samin Ghamami and Bo Zhang. Efficient monte carlo counterparty credit risk pricing and measurement. *Coleman Fung Risk Management Reasearch Center, Working Paper 2013 -02*, July 2013.

Paul Glasserman. *Monte Carlo Methods in Financial Engineering*. Applications of mathematics : stochastic modelling and applied probability. Springer, 2004. ISBN 9780387004518. URL http://books.google.cz/books?id=e9GWUsQkPNMC.

John Hull and Alan White. Pricing interest-rate-derivative securities. *Review of Financial Studies*, 3(4):573–92, 1990. URL http://ideas.repec.org/a/oup/rfinst/v3y1990i4p573-92.html.

The MathWorks Inc. Counterparty credit risk and cva. Financial Instruments Toolbox, 2014. URL http://www.mathworks.com/help/fininst/counterparty-credit-risk-and-cva.html.

Bank for International Settlements. Capital treatment for bilateral counterparty credit risk finalised by the basel commitee, June 2011. Available at http://www.bis.org/press/p110601.htm.

C. Kenyon and E. Stamm. *Discounting, Libor, CVA and Funding, Interest Rate and Credit Pricing*. Palgrave Macmillan, applied quantitive finance edition, 2012.

Michael Pykhtin and Steven Zhu. A guide to modeling counterparty credit risk. *A Guide to Modeling Counterparty Credit Risk*, July/August 2007. Available at SSRN: http://ssrn.com/abstract=1032522.

F. Vrins. Credit risk - getting cva up and running. *Risk : Managing Risk in the World's Financial Markets*, 2011.

John C. Williams and John B. Taylor. A black swan in the money market. *American Economic Journal: Macroeconomics*, 1(1):58–83, January 2009. URL http://ideas.repec.org/a/aea/aejmac/v1y2009i1p58-83.html.

Chapter 7
Content of Enclosed DVD

There is a DVD enclosed to this thesis which contains data and MatLab source codes.

- Folder 1: Source codes
- Folder 2: Data