Mathematics in Finance

UIMP-RSME Lluís A. Santaló Summer School
Mathematics in Finance and Insurance
July 16–20, 2007
Universidad Internacional Menéndez Pelayo
Santander, Spain

Santiago Carrillo Menéndez
José Luis Fernández Pérez
Editors
Mathematics in Finance
Mathematics in Finance

UIMP-RSME Lluis A. Santaló Summer School
Mathematics in Finance and Insurance
July 16–20, 2007
Universidad Internacional Menéndez Pelayo
Santander, Spain

Santiago Carrillo Menéndez
José Luis Fernández Pérez
Editors
Contents

Preface vii

Hedge Funds as Knock-Out Options
Marcos Escobar, Stefan Krämer, Florian Scheibl, Luis A. Seco and Rudi Zagst 1

Rough Paths based Numerical Algorithms in Computational Finance
Lajos Gergely Gyurkó and Terry Lyons 17

Hedge Funds
Luis A. Seco and Fangyuan Chen 47

Modeling and Pricing credit derivatives
Rudi Zagst and Matthias Scherer 111
This page intentionally left blank
Preface

This volume contains four selected survey papers on some aspects of mathematical finance presented and discussed at the “Lluis A. Santaló Summer School” that was held in Santander (Spain) on July 2007 as part of the activities of the Universidad Internacional Menéndez Pelayo (UIMP), in collaboration with the Real Sociedad Matemática Española (RSME).

The role of stochastic differential equations in finance is well known and so is the need for solving numerically these kinds of equations. In computational finance, path-wise approximation of solutions for stochastic differential equations is required, for example, when evaluating strongly path-dependent products, such as Asian options or range accrual products. The paper by L.G. Gyurko and T. Lyons presents a general framework for deriving high-order, stable and tractable path-wise approximations of Stratonovich stochastic differential equations.

Hedge funds have been one of the hot topics in finance in the last twelve years or more. At various times, some of them have been the focus of attention of the mass media. The paper by L. Seco and F. Chen presents a systematic survey on this topic from a mathematical point of view. The authors analyze, among other issues, the different classes of hedge funds and the models used both for management and risk control.

Still in the world of hedge funds, the paper by M. Escobar, S. Krämer, F. Scheibl, L. Seco and R. Zagst introduces a new theoretical framework to price hedge funds’ equity, inspired by the framework of Black and Cox for the valuation of companies’ equity as call options. The proposed approach is able to fit quite well the first four moments of the distribution of real returns when used with a sample of over a thousand hedge funds.

Credit derivatives have been among the most demanded products in the financial markets. The subprime crisis has placed this kind of product in the spotlight. R. Zagst and M. Scherer present a survey of the most widely used credit derivatives and analyze the most relevant approaches and models used in the financial sector: structural-default, intensity-based, reduced-form and hybrid models which combine the advantages of structural and intensity based models. They also focus on the modeling of joint defaults, one of the key issues for the pricing and risk measurement of this class of derivatives.

The editors wish to thank the Real Sociedad Matemática Española for giving them the opportunity to organize the Summer School. Our thanks also go to Universidad Internacional Menéndez Pelayo, a nice place with excellent organization and outstanding facilities, which made it easy for us to organize the lectures series within the School. Finally, the speakers at the Summer School and authors of this volume deserve our heartfelt gratitude both for the excellent lectures delivered at
the School and for their kind cooperation while preparing the corresponding notes which compose this volume.

Santiago Carrillo Menéndez and José Luis Fernández Pérez
Hedge Funds as Knock-Out Options

Marcos Escobar, Stefan Krämer, Florian Scheibl, Luis Seco, and Rudi Zagst

Abstract. This paper introduces a new theoretical framework to price hedge funds’ equity. It is inspired on the famous framework of Black and Cox for the valuation of companies’ equity as call options. Our structural model describing hedge funds uses barrier options (i.e. down-and-out call options as well as up-and-out put options) to allow for the special structure of hedge funds’ debt position. The quality of these models is evaluated by its capability to reproduce a high range of historical hedge fund returns. Different variations of the model are compared using this criteria. To fit the model parameter to real hedge fund data the method of moments is used. The application of the models to a set of over 1000 hedge funds showed that the model fits the first four moments of the real returns quite well. Especially the documented stylized features of high kurtosis and skewness are very well captured by this model.

1. Introduction

Although the beginnings of hedge funds are now more than 50 years ago they had their breakthrough in the last decade. Especially big crashes like that of famous LTCM led to more public attention on hedge funds.\(^1\) Again, in the subprime crisis in the summer of 2007 hedge funds got return to the public eye: Some for gaining tremendous returns because they anticipated the crisis, others for being closed because of bankruptcy.\(^2\)

It is very difficult to explain the past returns of hedge funds because of two reasons, first a lack of information about the company’s portfolio, secondly the sparsity of data. This is the nature of hedge funds. The strategy is the core asset of a hedge fund. Often anomalies in the capital markets or opportunities for statistical arbitrage are exploited. Therefore it is obvious that predicting the behaviour of hedge funds is an even more challenging task in the future. Most models try to explain hedge fund returns using very sophisticated and complex approaches. For example, several techniques have been developed to identify the structure of the portfolio implicitly by observing the historical returns of the funds. Needless to say this approach is doom in most cases because of lack of information.

\(^1\)cp. Edwards (1999)
A popular alternative are factor models, here the return of a hedge fund is explained by several observable indicators like market indices (cp. Fung (1997), Agarwal (2000)). Some of these models use a multitude of factors, for example stock indices, bond spreads, put and call options or commodity futures. This approach immediately raises various questions, for instance, how these factors should be chosen, e.g. what strike price for an option should be used or which specific indices should be considered.

The purpose of this paper is not to come up with another factor model with even more factors and complexity than existing models. In fact we use a different approach and try to develop a much simpler model which does not need a lot of market data but also explains the return of hedge funds quite well. We place emphasis on the special capital structure of hedge funds which is often characterized by a high level of leverage. Therefore we use the traditional option model of Black-Scholes (1973) and the extension of this model towards barrier options by Black-Cox (1976) as starting point of our consideration. In our framework, the value of a hedge fund is determined by the value of its assets and the face value of debt.

Factor models also use options to explain hedge fund returns, but they only use them to describe a single position of the hedge funds’ assets (see ...). Whereas in this paper we use knock-out options as framework for a structural model. Therefore a default before maturity is possible, which is implicitly ruled out in the model of Merton 1973 which is widely used to valuate traditional companies. In the context of hedge funds this point of view seems to be more realistic, because there are significant differences between the characteristics of industrial enterprises and hedge funds. Whereas bonds and bank loans with a certain maturity are the most important part of a companies’ debt, this is not true for hedge funds as well. Although the concept of leverage is very important for the hedge fund industry, the classic forms of debt are not the only thing to look at.\(^3\)

The economic value of debt also consists of off-balance-sheet leverage through the usage of derivatives. In most industry enterprises the leverage effect of positions in derivatives is negligibly small compared to the total volume of debt, because the usage of these instruments is often limited to hedge certain eminent business risks. In the hedge fund industry using derivatives can be seen as inherent part of the business concept. So not taking into account this type of leverage would obviously lead to wrong results.

Another characteristic of hedge funds is their dependency of a so called „prime broker“ (specified by the foundation contract of the fund). This prime broker administrates all the assets and liabilities of the hedge fund. At the same time this broker fulfills a set of other duties as well. He settles the deals, provides account statements and financing for leverage. Therefore the prime broker is well informed about the status of the fund. If the hedge fund’s assets will do bad, he is one of the first partners in the market to recognize this.\(^4\)

Another fact is that many of the hedge fund’s liabilities are collected on the margin account at a prime broker. Since margin accounts are adjusted daily, there is not a specified maturity of this debt position. This altogether leads to the assumption, that a prime broker is able to default a hedge fund easily when things

\(^3\)cp. FMA (2005), pp. 62 et seq.
HEDGE FUNDS AS KNOCK-OUT OPTIONS

...go bad. That is why the implicit assumption of no default before maturity in the model of Black and Scholes is not acceptable for hedge funds at all.

In the next section we develop two barrier option models for hedge funds. In the third section we apply this model on historical hedge fund data and analyze its behavior. In the last section we summarize the results and draw conclusions.

2. The Model

2.1. Down-and-Out Call Model. In this section, the notion of a Hedge Fund’s equity as a “Down and Out Option” on its assets is studied. This approach has solid financial ground (see Black 1976 and Rubinstein et al. 1991) for standard companies, therefore it is a natural step in the understanding of HFs. The nature of the equity of a HF is more complex than that of other companies as it could be a portfolio with a changing mixture of a wider type of positions. Hence this approach handles the complexity on the assets and the debt instead.

We start this section presenting some general assumptions that will be used in this paper. To the model using Down-and-Out Call (DOC) options might also be referred to as basic model while the Up-and-Out Put (UOP) might be called the alternative model.

Assumption 1 (underlying process): The asset value $V_t$ of a hedge fund is a log-normally distributed random variable with the dynamic $dV_t := \mu_V dt + \sigma_V dW_t$ whereas $\mu_V$ and $\sigma_V$ are constant and $W_t$ stands for a standard Brownian Motion process.

This means $X_t = \ln(V_t)$ is normally distributed with mean $(\mu_V - \frac{1}{2}\sigma_V^2)t$ and standard deviation $\sigma_V \sqrt{t}$:

$$dX_t = \mu_X dt + \sigma_V dW_t \quad \text{with} \quad \mu_X = \mu_V - \frac{\sigma_V^2}{2}$$

For simplification we always set $V_0 = 1$ which is equivalent to $X_0 = 0$.

Assumption 2 (payouts to equity holders). The only payment from the hedge fund to the equity holders is at maturity $T$. The hedge fund’s assets are not due to changes caused by additional or reduced investments or dividend payments.

Assumption 3 (market properties). The market is assumed to be arbitrage-free and complete. There is a risk-free asset delivering a constant interest rate $r$.\footnote{We assume a riskless interest rate $r = 0.07$ which has been the average historical rate of return of German government bonds with a maturity of 10 years in the period from 1955 to 2003, cp. Stehle (2004)}

From this assumption we know that the Martingal Equation must hold for the hedge fund’s asset under the risk-neutral probability measure:

$$(2.1) \quad V_t = e^{-r(T-t)} \cdot E[V_T | F_t]$$

$F_t$ stands for the filtration up to time $t$ associated to the process $V_t$. The most simple and intuitive way to build a hedge fund model with knock-out is a model with a constant barrier defining the knock-outs. Therefore we introduce the following assumption:
Assumption 4. The hedge fund has a constant amount of debt $D$. If the value of the hedge fund falls below this amount before maturity, then the cash flow to equity holders will be zero. If this barrier is not touched, the final payment to equity is $V_T - D$.  

This means that the hedge fund’s equity $S_t$ is a Down-and-Out Call Option on the value $V_t$ where the constant Debt $D$ equals both the strike and the barrier of the option. This can be written in the following way:

$$S_t = e^{-r(T-t)}E[(V_T - D) \cdot 1_{\tau > T} | F_t]$$

with $\tau := \inf\{s | s \geq 0, V_s \leq D\}$.

Some possible movements of the asset value are illustrated in Figure 2.1 on the left side. Paths that touch the barrier before maturity are marked by a dotted line. It should be noted that there is a positive probability for recovering of a “knocked out” path in the world of a standard Black-Scholes Call-Option. This probability is cut off in the environment of a knock-out option. The figure on the right side shows the density of the endpoint of a log-normally distributed asset and the endpoint’s density under a no-knock-out condition. Note the distribution of knock-outs is a mixing of a discrete and continuous, therefore there is a mass at $D$ complementing the density in the figure.

Figure 2.1. Movement of asset value (left picture) and densities of the endpoint’s distribution with or without knock-outs (right picture)
There has been a lot of research about knock-out options in the last decades.\footnote{cp. Merton (1973), p. 175 et seq.; cp. Black (1976), Valuing Corporate Securities, p. 356; cp. Chesney (1997), p. 5 et seq.; cp. Reiner (1991)} Their risk-neutral valuation is possible with a closed-form solution that was already found by Rubinstein et. al (1991). The option value is that of a plain-vanilla European Black-Scholes call option $S_t^{Call}$ reduced by a “knock-out discount” (KOD):

\begin{equation}
S_t^{DOC}(V_t, D) = S_t^{Call}(V_t, D) - KOD(V_t, D)
\end{equation}

with

\begin{equation}
KOD(V_t, D) = (\frac{D}{V_t})^{\frac{\sigma^2}{2}} (V_t \cdot \Phi(c_1) - D \cdot e^{-r\tau} \cdot \Phi(c_2))
\end{equation}

\begin{equation}
c_{1,2} = \frac{ln \left( \frac{D}{V_t} \right) + (r \pm \frac{1}{2} \sigma^2)\tau}{\sigma \sqrt{\tau}}
\end{equation}

The analysis of this kind of option is very interesting: Its development is very likely to be almost a linear transformation of the underlying. This behavior is observable when the value of the underlying is far above the strike and barrier of the option. Because we introduced the model to explain non-Gaussian returns, we will concentrate on the case with a high debt ratio. For high values of $D$ the option behavior becomes non-linear compared to the underlying. This can be seen in figure 2.2 which shows the derivative of the option value with respect to the value of the underlying, the so-called option delta. In a region where this option delta is very close to one, the behavior of the option is linear on the behavior of the underlying.
2.2. Reachable Moments with the DOC Model. The goal of using an option-based structure to model hedge-funds is to create returns that have similar properties as real hedge fund returns. To describe the behavior of such returns a meaningful way to describe the behavior of Hedge funds’ returns is to look at their central moments: mean, volatility, skewness and kurtosis. In order to analyze how flexible our model is and to check how well it can rebuild the behavior of real hedge funds, we determine the ranges and the possible combinations of moments that this model can produce.
Mean and volatility can be almost perfectly fitted for relevant values which have been observed in real hedge fund data.\footnote{This is basically achieved by adjusting the parameters $\mu_V \quad \text{and} \quad \sigma_V$. For details see section 3.} But we need to analyze the possible values for skewness and kurtosis. Figure 2.3 shows the maximal reachable values for these two moments dependent to mean and volatility. While the influence of the standard deviation is rather unimportant the value of the mean has a significant impact. For positive mean we can reach very high skewness above five while for negative mean the maximum skewness is zero. For the kurtosis we are able to reproduce very high

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{skewness_kurtosis.png}
\caption{Maximum reachable values for skewness and kurtosis}
\end{figure}
values above 20 for the full range of mean, which is fully sufficient for nearly all hedge funds.

On the other hand also the minimal reachable values are important. Because the volatility has few influence we look at two-dimensional plots of possible moment combinations in Figure 2.4. On the left side we see that it is impossible to get negative mean and positive skewness or positive mean and skewness values below -0.1. The latter is the quite more relevant problem because most hedge funds have positive mean and many are also negatively skewed. The second model restriction can be seen on the right side: It is not possible to generate kurtosis below three which accords to the fact that all long options have high kurtosis. Especially the restriction to positive skewness is problematic because most hedge fund returns have positive mean and negative skewness. Therefore we will discuss a slightly different model to enlarge the range of possible moment combinations.

2.3. Up-and-Out Put Model. Hedge Funds use a variety of investment methods, one of them is short selling. This behaviour used to be the principal investment approach for some HF’s companies during fixed periods of time. For those cases, the equity value could be seen as the expected value of a portfolio of short selling positions. Therefore we summarize this portfolio as a single stochastic debt process. The collateralization of this portfolio is conducted via a trust consisting of cash or treasury bonds.

Assumption 1 (underlying process): The liability value $L_t$ of a hedge fund is a log-normally distributed random variable with the dynamic

$$\frac{dL_t}{L_t} := \mu_L dt + \sigma_L dW_t$$

whereas $\mu_L$ and $\sigma_L$ are constant.

The assumptions two (only one payout to equity at maturity) and three (constant riskless interest rate) remain untouched. Whereas the payout structure is defined in accordance with the different definition of the cash amount:

Assumption 4. The hedge fund has a growing amount of cash $C_t = C_0 \cdot e^{\alpha t}$. If the liabilities of the hedge fund $L_t$ exceed this amount before maturity, then the cash flow to equity holders will be zero. If this barrier is not touched, the final payment to equity is $C_T - L_T$.

---

The structure of this option is just the opposite of the Down-and-Out Call with the difference of a growing barrier $C_t$. The valuation of such an Up-and-Out Put option is also very similar. It consists of a standard Black-Scholes Put which is reduced by another knock-out-discount (KOD):

\[
S_t^{UOP}(L_t, C_t, \alpha) = S_t^{Put}(L_t, C_t, \alpha) - KOD(L_t, C_t, \alpha)
\]

with

\[
KOD(L_t, C_t, \alpha) = -(\frac{C_t}{L_t})^{2(r-\alpha)} \left( C_t \cdot \Phi(-c_1) - L_t \cdot e^{-(r-\alpha)\tau} \cdot \Phi(-c_2) \right)
\]

\[
c_{1,2} = \frac{\ln \left( \frac{C_t}{L_t e^{\alpha\tau}} \right) + (r \pm \frac{1}{2} \sigma^2 L)\tau}{\sigma L \sqrt{\tau}}
\]

where $\alpha$ is the growth rate of $C$. 10

2.4. Reachable Moments with the UOP Model. In analogy to section 2.2 we analyze the attainability of moments for the alternative model. Figure 2.5 shows that it is possible to get negative skewness down to -0.3 with small positive values of the mean. The enlargement of the space of reachable moments is not very big, but lots of hedge funds lie in exactly this area and so this is an important improvement to the model before.

By means of using a value for $\alpha$, where $\alpha > r$ it would be possible to achieve even more extreme negative skewness (this a mathematical trick to achieve a wider range of behaviours for the equity.). In order to keep this model assumption reasonable and to reduce complexity in fitting we just use this one parameter value.

![Figure 2.5. Possible values for mean/skewness with the UOP model](image)

3. Application on real data

In the previous sections we presented two different models to describe a hedge fund as knock-out option. The quality of these models has been indicated by their capability to create a high range of moments and their rather small complexity.
In the following section the empirical foundation of our model is fostered. So we compare the returns created by our model with historical data of real hedge funds. We were very keen about the outcome of this section, because our models had to prove their real world relevance for the first time. An important step for this is the parameter fitting. For every fund we need to calibrate the model parameters so they best fit the data. The goal is to provide a "model hedge fund" that behaves very similar to the real fund.

3.1. Description of the data. There are no standardized publication rules for hedge funds and therefore, most of them report only few data. Some hedge funds provide daily return data but most of them only report monthly returns. Furthermore, many hedge funds are very recent (less than 5 years of history, around 60 data points). If we restrict our analysis on funds which have existed for more than 5 years, we observe at least 60 datapoints. Many funds report also their Net Asset Values every month but this might be a misleading figure because it can also change as a result of investment inflows and outflows to the equity holders.

In this paper we work with 1529 funds of 15 different strategies from the Wharton Hedge Fund database. The average moments for each style\footnote{Unweighted averages; mean and standard deviation are annualized values whereas skewness and kurtosis are based on the whole 60 months.} can be found in table 1. Here the moments were computed assuming time independent equity returns. Most strategies have negative skewness and positive excess kurtosis but hedge funds are very heterogeneous and there are funds with almost every possible combination of moments.

<table>
<thead>
<tr>
<th>Strategy</th>
<th>#</th>
<th>Mean</th>
<th>Std</th>
<th>Skew</th>
<th>Kurt</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>1529</td>
<td>8.82%</td>
<td>11.50%</td>
<td>-0.1664</td>
<td>4.7056</td>
</tr>
<tr>
<td>FOF-Multi Strategy</td>
<td>362</td>
<td>7.23%</td>
<td>4.92%</td>
<td>-0.3226</td>
<td>4.4798</td>
</tr>
<tr>
<td>Equity Long/Short</td>
<td>264</td>
<td>9.90%</td>
<td>13.94%</td>
<td>-0.1390</td>
<td>4.5823</td>
</tr>
<tr>
<td>N/A</td>
<td>227</td>
<td>6.66%</td>
<td>14.69%</td>
<td>-0.0783</td>
<td>4.0113</td>
</tr>
<tr>
<td>CTA-Systematic/Trend-Following</td>
<td>174</td>
<td>6.77%</td>
<td>18.16%</td>
<td>0.0351</td>
<td>3.9215</td>
</tr>
<tr>
<td>Emerging Markets</td>
<td>84</td>
<td>22.14%</td>
<td>16.97%</td>
<td>-0.1784</td>
<td>4.5728</td>
</tr>
<tr>
<td>Sector</td>
<td>70</td>
<td>9.31%</td>
<td>16.63%</td>
<td>-0.1664</td>
<td>4.6874</td>
</tr>
<tr>
<td>Convertible Arbitrage</td>
<td>49</td>
<td>8.93%</td>
<td>7.03%</td>
<td>0.1123</td>
<td>4.8761</td>
</tr>
<tr>
<td>Event Driven Multi Strategy</td>
<td>49</td>
<td>10.84%</td>
<td>10.06%</td>
<td>-0.4101</td>
<td>5.5077</td>
</tr>
<tr>
<td>FOF-Single Strategy</td>
<td>47</td>
<td>7.07%</td>
<td>6.50%</td>
<td>-0.3735</td>
<td>5.5674</td>
</tr>
<tr>
<td>Equity Market Neutral</td>
<td>42</td>
<td>5.56%</td>
<td>9.12%</td>
<td>-0.4114</td>
<td>6.4443</td>
</tr>
<tr>
<td>CTA-Discretionary</td>
<td>35</td>
<td>9.75%</td>
<td>15.58%</td>
<td>-0.0136</td>
<td>6.6213</td>
</tr>
<tr>
<td>Global Macro</td>
<td>34</td>
<td>11.41%</td>
<td>13.84%</td>
<td>-0.0231</td>
<td>4.8701</td>
</tr>
<tr>
<td>Relative Value Multi Strategy</td>
<td>31</td>
<td>7.23%</td>
<td>3.26%</td>
<td>0.3118</td>
<td>3.1215</td>
</tr>
<tr>
<td>Merger Arbitrage</td>
<td>31</td>
<td>11.65%</td>
<td>4.17%</td>
<td>0.2129</td>
<td>1.5179</td>
</tr>
<tr>
<td>Distressed Securities</td>
<td>30</td>
<td>10.27%</td>
<td>5.81%</td>
<td>0.8168</td>
<td>2.9753</td>
</tr>
</tbody>
</table>

Table 1. Hedge Fund data overview

3.2. The estimation problem. The moments we can reach with the two alternative models are different and we decide which to use from the hedge fund’s
skewness. For funds with positive skewness we use the DOC model and for those with negative skewness the UOP model.

To provide comparable information we always standardize \( V_0 = 1 \) respectively \( L_0 = 1 \). Then we can calculate the leverage factor in \( t = 0 \) by \( \frac{D}{1-D} \) respectively \( \frac{1}{C_0-1} \).

The standardization is possible because the valuation formulas are linear in the following way:

\[
S_t^{DOC}(kV_t,kD) = k \cdot S_t^{DOC}(V_t,D)
\]

\[
S_t^{UOP}(kL_t,kC_t,\alpha) = k \cdot S_t^{UOP}(L_t,C_t,\alpha)
\]

We can calculate these option values with formula 2.3 and formula 2.4 from the path of \( V_t \) respectively \( L_t \). Because these are stochastic processes we need to simulate many paths to get the moments of \( S_t \). To simulate and calculate \( S_t \) we need the parameters \( r, \tau, D, \sigma_V, \mu_V \) for the DOC model or \( r, \tau, C_0, \sigma_L, \mu_L \) for the UOP model.

The value of \( r \) as riskless interest rate should be independent from the regarded hedge fund and so we do not fit it to the data of a single fund. \(^{12}\)

For the time to maturity \( \tau \) we use ten years because we think this is a good estimation for the time horizon of hedge fund managers. \(^{13}\) But most hedge funds do not exist for such a long time: Between 1989 and 1996 the average life time of hedge funds was only 40 months. \(^{14}\) A newer study shows that the attrition rate of hedge funds is about 7 to 9 percent per year. \(^{15}\) This implies an average life time of about 11 to 14 years. So the ten years we assumed for the time to maturity of a fund is right in the middle of the values the two studies propose.

Hence, we concentrate on estimating the unknown parameters \( D, \sigma_V, \mu_V \) respectively \( C_0, \sigma_L, \mu_L \).

### 3.3. Parameter fitting with the method of moments.

To estimate the missing parameters we use the method of moments (MoM). The application of MoM requires at least as many conditions (moments) as parameters which should be fitted through this approach. We use the first four central moments mean, standard deviation, skewness and kurtosis, all calculated from the step-to-step returns \( R_t \) of \( S_t \).

We use the standard methods to estimate the moments of a historical path. With four moments for three fitted parameters we are able to create a sounding estimation framework. At first we need the moments \( (m_1, m_2, m_3, m_4) \) of the historical returns \( R_t \). These will be compared to the moments \( (\hat{m}_1, \hat{m}_2, \hat{m}_3, \hat{m}_4) \) from our model for a specific triple of parameters. These moments are obtained by simulating several paths for \( V_t \) and calculate \( \hat{S}_t \). For each path we calculate the moments and then average over all paths which have not been knocked out during the observed period.

\[
\hat{m}_i := E \left[ m_i|\hat{D}, \hat{\mu}_V, \hat{\sigma}_V \right] \text{ respectively } \hat{m}_i := E \left[ m_i|\hat{C}_0, \hat{\mu}_L, \hat{\sigma}_L \right], i = 1, 2, 3, 4
\]

\(^{12}\) In accordance to assumption 3.

\(^{13}\) We keep the time to maturity constant to avoid reduction in hedge fund value only due to the spend of time.

\(^{14}\) cp. Edwards (1999), p. 192

\(^{15}\) cp. Baba (2006), p. 30
Now we compare the moments from the model with those of the historical returns $R_t$. We find the best parameter estimation by least square optimization with criterion:

$$
\min_{D, \mu_V, \sigma_V} \sum_{i=1}^{4} [w_i(m_i - \hat{m}_i)]^2 \text{ respectively } \min_{C, \tilde{\mu}_L, \tilde{\sigma}_L} \sum_{i=1}^{4} [w_i(m_i - \tilde{m}_i)]^2
$$

To regard the different levels of the absolute values the four central moments can reach we use a vector of weighting factors $w_i$. Especially the kurtosis reaches much higher values than the other moments and would dominate the fitting without such a weighting. Therefore it has proven to be reasonable to use $w = (10, 10, 10, 1)$. Moreover, the kurtosis is difficult to obtain precisely (for small sample sizes) so it should be given less importance.

In practice the calculation of moments with the model needs a lot of computational power and time. In order to lessen this problem we use a set of standard parameter values and build up a grid of moments for all their combinations. To find the best set of parameters for a fund we compare its moments to every entry in this grid. In order to get a closer meshed grid it is also useful to use linear interpolation in the spaces between the simulated results. In a second step this fitting can be further improved by using the computed parameters as starting point for a least-square optimization algorithm.

3.4. Results. At first we look at the fitting outcome with the DOC model. We analyze how far distant the real moments and those generated by the model are in order to measure the quality of the fitting. A good fitting means the model fits the fund better than [1% 2% 2% 20%] for the four moments mean, standard deviation, skewness and kurtosis.$^{16}$

<table>
<thead>
<tr>
<th>Sector</th>
<th>#pos. skew</th>
<th>#good</th>
<th>Leverage</th>
<th>$\sigma_V$</th>
<th>$\mu_V$</th>
</tr>
</thead>
<tbody>
<tr>
<td>FOF-Multi Strategy</td>
<td>79</td>
<td>75</td>
<td>14.95</td>
<td>1.88%</td>
<td>7.54%</td>
</tr>
<tr>
<td>Equity Long/Short</td>
<td>106</td>
<td>96</td>
<td>7.55</td>
<td>4.13%</td>
<td>10.28%</td>
</tr>
<tr>
<td>N/A</td>
<td>92</td>
<td>85</td>
<td>3.37</td>
<td>6.45%</td>
<td>8.02%</td>
</tr>
<tr>
<td>CTA-Systematic/Trend-Following</td>
<td>89</td>
<td>73</td>
<td>4.25</td>
<td>5.14%</td>
<td>8.27%</td>
</tr>
<tr>
<td>Emerging Markets</td>
<td>32</td>
<td>24</td>
<td>9.00</td>
<td>4.56%</td>
<td>14.90%</td>
</tr>
<tr>
<td>Sector</td>
<td>26</td>
<td>22</td>
<td>11.33</td>
<td>3.48%</td>
<td>8.57%</td>
</tr>
<tr>
<td>Convertible Arbitrage</td>
<td>23</td>
<td>21</td>
<td>25.88</td>
<td>2.11%</td>
<td>9.95%</td>
</tr>
<tr>
<td>Event Driven Multi Strategy</td>
<td>9</td>
<td>7</td>
<td>7.60</td>
<td>5.37%</td>
<td>13.71%</td>
</tr>
<tr>
<td>FOF-Single Strategy</td>
<td>14</td>
<td>14</td>
<td>17.52</td>
<td>1.09%</td>
<td>6.89%</td>
</tr>
<tr>
<td>Equity Market Neutral</td>
<td>14</td>
<td>11</td>
<td>32.90</td>
<td>2.45%</td>
<td>8.09%</td>
</tr>
<tr>
<td>CTA-Discretionary</td>
<td>20</td>
<td>18</td>
<td>12.53</td>
<td>4.58%</td>
<td>11.64%</td>
</tr>
<tr>
<td>Global Macro</td>
<td>18</td>
<td>18</td>
<td>16.45</td>
<td>1.86%</td>
<td>9.75%</td>
</tr>
<tr>
<td>Relative Value Multi Strategy</td>
<td>10</td>
<td>9</td>
<td>150.52</td>
<td>1.43%</td>
<td>6.61%</td>
</tr>
<tr>
<td>Merger Arbitrage</td>
<td>15</td>
<td>15</td>
<td>102.09</td>
<td>1.13%</td>
<td>5.13%</td>
</tr>
<tr>
<td>Distressed Securities</td>
<td>20</td>
<td>20</td>
<td>66.11</td>
<td>2.19%</td>
<td>10.28%</td>
</tr>
</tbody>
</table>

Table 2. DOC Fitting Results

As we can see, we estimate very different leverage values for different strategies, but all levels are rather high. These values represent the economic leverage of a

$^{16}$Skewness and kurtosis again refer to annualized values.
fund which depends not only on the amount of real debt but is also very likely influenced by derivative business that also lever the returns.

The results of the negative skewed funds using the UOP model show a surprisingly low initial leverage. Whereas the variance is rather small and the mean of the short asset is higher than the risk less rate $r$.

<table>
<thead>
<tr>
<th>Strategy</th>
<th>#neg. skew</th>
<th>#good</th>
<th>Leverage</th>
<th>$\sigma_L$</th>
<th>$\mu_L$</th>
</tr>
</thead>
<tbody>
<tr>
<td>FOF-Multi Strategy</td>
<td>283</td>
<td>205</td>
<td>1.21</td>
<td>1.27%</td>
<td>12.08%</td>
</tr>
<tr>
<td>Equity Long/Short</td>
<td>158</td>
<td>86</td>
<td>1.91</td>
<td>2.64%</td>
<td>11.47%</td>
</tr>
<tr>
<td>N/A</td>
<td>135</td>
<td>105</td>
<td>2.70</td>
<td>2.88%</td>
<td>10.56%</td>
</tr>
<tr>
<td>CTA-Systematic/Trend-Following</td>
<td>85</td>
<td>59</td>
<td>3.25</td>
<td>4.05%</td>
<td>10.05%</td>
</tr>
<tr>
<td>Emerging Markets</td>
<td>52</td>
<td>14</td>
<td>2.53</td>
<td>3.50%</td>
<td>9.50%</td>
</tr>
<tr>
<td>Sector</td>
<td>44</td>
<td>18</td>
<td>2.02</td>
<td>3.56%</td>
<td>10.72%</td>
</tr>
<tr>
<td>Convertible Arbitrage</td>
<td>26</td>
<td>22</td>
<td>1.47</td>
<td>3.27%</td>
<td>11.77%</td>
</tr>
<tr>
<td>Event Driven Multi Strategy</td>
<td>40</td>
<td>16</td>
<td>1.07</td>
<td>3.88%</td>
<td>12.31%</td>
</tr>
<tr>
<td>FOF-Single Strategy</td>
<td>33</td>
<td>22</td>
<td>1.47</td>
<td>1.40%</td>
<td>11.86%</td>
</tr>
<tr>
<td>Equity Market Neutral</td>
<td>28</td>
<td>21</td>
<td>1.34</td>
<td>1.76%</td>
<td>12.52%</td>
</tr>
<tr>
<td>CTA-Discretionary</td>
<td>15</td>
<td>8</td>
<td>1.16</td>
<td>2.13%</td>
<td>12.00%</td>
</tr>
<tr>
<td>Global Macro</td>
<td>16</td>
<td>6</td>
<td>2.86</td>
<td>2.33%</td>
<td>10.33%</td>
</tr>
<tr>
<td>Relative Value Multi Strategy</td>
<td>22</td>
<td>11</td>
<td>1.25</td>
<td>1.45%</td>
<td>12.27%</td>
</tr>
<tr>
<td>Merger Arbitrage</td>
<td>18</td>
<td>10</td>
<td>1.19</td>
<td>1.10%</td>
<td>12.20%</td>
</tr>
<tr>
<td>Distressed Securities</td>
<td>10</td>
<td>2</td>
<td>1.00</td>
<td>1.00%</td>
<td>11.00%</td>
</tr>
</tbody>
</table>

Table 3. UOP Fitting Results

Table 2 shows the UOP framework achieve a good fitting for more than 50% of the negatively skewed HF, this result is quite consistent among all HF strategies except for “Equity Market Neutral” and “Convertible Arbitrage”. In comparison to the DOC approach, the percentage of good fitting is lower. These results are expected as the DOC framework (assets stochastic and constant debt) is more compatible with the way most companies work (more acceptable in the financial literature) than the UOP (constant assets, stochastic debt).

Levels of hedge fund leverage by strategy can be found in the literature.\cite{17}

<table>
<thead>
<tr>
<th>strategy</th>
<th>leverage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fixed Income Arb.</td>
<td>19</td>
</tr>
<tr>
<td>Other long short</td>
<td>4</td>
</tr>
<tr>
<td>Long Only</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 4. Implied leverage by strategy

4. Conclusion

Conclusion. This work introduces a new theoretical framework to price hedge funds’ equity based on the structural framework of Black-Cox (1976) for the valuation of companies’ equity as call barrier options (i.e. down-and-out call options as well as up-and-out put options). The quality of these models is evaluated by

\cite{17} cp. Blundell-Wignall (2007), p. 48
its capability to reproduce a high range of historical hedge fund returns. Different variations of the model are compared using this criteria. The method of moments is used to fit the model parameter to real hedge fund data. The application of the models to a set of over 1000 hedge funds showed that the model fits the first four moments of the real returns quite well. Especially the documented stylized features of high kurtosis and skewness are very well captured by this model.

References


H. M. Kat, W. Lu (2002): 
‘An Excursion into the Statistical Properties of Hedge Funds’, ISMA Discussion 
Papers in Finance 2002-12.

R. Merton (1973):
'Theory of Rational Option Pricing', Bell Journal of Economics and Management 

‘Breaking Down the Barriers’, RISK 4, pp. 28-35.

R. Stehle (2004):
'Die Festlegung der Risikoprämie von Aktien im Rahmen der Schätzung des Wertes 
926.

M. Teo (2007):
'Lessons from the Sub-Prime Crisis', Statistical Digest (SMU).

DEPARTMENT OF MATHEMATICS, RYERSON UNIVERSITY, 350 VICTORIA ST., TORONTO, ON-
TARIO M5B 2K3, CANADA
E-mail address: escobar@ryerson.ca

FIM PROGRAM, TU MÜNCHEN, BOLTZMANNSTR. 3, 85748 GARCHING, GERMANY
E-mail address: pfefferminze@googlemail.com

FIM PROGRAM, TU MÜNCHEN, BOLTZMANNSTR. 3, 85748 GARCHING, GERMANY
E-mail address: flo.scheibl@googlemail.com

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF TORONTO, TORONTO, ONTARIO M5S2E4, 
CANADA.
E-mail address: seco@math.utoronto.ca

HVB-STIFTUNGSINSTITUT, TU MÜNCHEN, BOLTZMANNSTR. 3, 85748 GARCHING, GERMANY
E-mail address: zagst@ma.tum.de
This page intentionally left blank
Rough Paths based Numerical Algorithms in Computational Finance

Lajos Gergely Gyurkó and Terry Lyons

Abstract. The paper connects asymptotic estimations of \([4]\) and \([8]\) with the Rough Paths perspective (\([14]\), \([15]\)) to present a general framework for deriving high order, stable and tractable path-wise approximations of stochastic differential equations. The approach, which can be traced back to \([19]\) and probably earlier, is based on locally deriving and solving random ordinary differential equations. A sufficient condition on the accuracy of the numerical ODE solver is given to ensure the global order is \(\gamma - 1/2\) if the local order is \(\gamma\). We also point out some practical solutions which make the high order schemes tractable.

1. Introduction

This paper describes methods for approximating solutions to Stratonovich differential equations in \(\mathbb{R}^n\) of the type

\[
\begin{align*}
    d\xi_t &= V_0(\xi_t)dt + \sum_{i=1}^d V_i(\xi_t) \circ dB^i_t \\
    \xi_0 &= x_0
\end{align*}
\]

where \(B_t = (B^1_t, \ldots, B^d_t)^T\) denotes a \(d\)-dimensional Brownian motion and the coefficient functions \(V_i, i = 0, \ldots, d\) are smooth \(\mathbb{R}^n \to \mathbb{R}^n\) functions with bounded derivatives up to a certain order. Throughout the paper we assume that the coefficient functions \(V_i\) for \(i = 0, \ldots, d\) satisfy sufficient regularity conditions implying the existence and uniqueness of solution to the SDE.

The central aim is to derive high order strong approximations of \(\xi_T\). We will say, that an approximation \(\hat{X}_T\) of \(\xi_T\) has (global) strong order \(\gamma\) if there exists a constant \(C\), not depending on the number of time steps and a positive \(\varepsilon\) such that for all \(\Delta t \in (0, \varepsilon)\)

\[
\mathbb{E} \left[ \left\| \xi_T - \hat{X}_T \right\|^2 \right]^{1/2} < C \Delta t^\gamma
\]

1991 Mathematics Subject Classification. Primary 60H35, 60H99, 65C30.
Key words and phrases. Stochastic Differential Equations.
The first author was supported in part by EPSRC scholarship EP/P501709/1.
Both authors were supported in part by the Oxford-Man Institute of Quantitative Finance.

©2010 American Mathematical Society
for some norm $\| \cdot \|$ on $\mathbb{R}^n$ where $\Delta t$ denotes the length of the longest time step produced by the discretization.

The numerical methods - derived in the paper - are based on an approach which has been established under the framework of the Rough Paths perspective ([14], [15]), albeit some of the key ideas can be traced back to earlier (see references in [15]). Although stochastic differential equations satisfying certain fairly general conditions can be interpreted as path-wise differential equations driven by rough paths ([15]), in this paper we stick to the use of the traditional Itô and Stratonovich stochastic integrals and SDEs while extensively applying some of the key ingredients of Rough Paths Theory, in particular: the representation of paths by their signature and log-signature. The paper highlights the importance of the relation between stochastic-Taylor expansion of solutions to SDEs and the signature of the Brownian paths driving the SDE. This relation is exploited when we derive random ordinary differential equations (ODEs) locally with solutions close to the solution of the SDE. In the paper, we emphasize the importance and efficiency of the numerical approximations of solutions to SDEs based on the (numerical) solution of these random ODEs.

In finance, the non-anticipating Itô integrals appear naturally. Nevertheless, our reason behind working with Stratonovich integrals is purely technical (the Stratonovich Brownian signature has simpler form than the Itô signature) and it is more precise to say that we are working with the Stratonovich form of Itô integrals. The principles of transforming Itô integrals to Stratonovich ones are recalled in section 2.3.

In computational finance, path-wise approximation of solutions for SDEs is required for example when evaluating path-dependent products (Asian options, Bermudan options, knock-out options etc.). However, note that some simple versions of these problems can be transformed to a European-type structure, and strong approximation is no longer necessary.

The paper is structured as follows. In section 2, the notation is introduced and some basic lemmas are reviewed. In section 3, we recall some ingredients of Rough Paths Theory applied throughout the paper and then in Theorem 3.1 we rephrase and prove the main result of [4] in terms of the Rough Paths terminology. The random ordinary differential equations are derived from the log-signature theorem of section 3. In section 4, we give a general sufficient condition (Theorem 4.1) for the global convergence of the numerical schemes as well as showing (Corollaries 4.1 and 4.2) when the Theorem 3.1 based schemes satisfy this condition. We derive how accurately one has to approximate the solution to the derived ODEs to ensure the high order global convergence. The high order schemes require the simulation of high order Brownian iterated integrals, however one can make small enough error without loosing the global convergence’s high order. In section 5, we show how one can approximate the Brownian iterated integrals based on piece-wise linear interpolations of the Brownian paths on a fine sub-scale. A sufficient condition on the fineness of the sub-scale preserving the scheme’s order of convergence is provided by Theorem 5.1. Finally in section 6, we demonstrate the efficiency of the log-signature theorem based approximations for some numerical examples.
2. Notation and some technicalities

In the following subsections, we introduce some rather technical but necessary notation and assertions. Besides the new notation, we focus on deriving $L^2$-bounds on Stratonovich iterated integrals. These bounds are important for deriving local and global error estimates of the numerical schemes.

2.1. Multi-indices. In the present paper, we introduce objects indexed by multi-indices of the form

\[ J = (j_1, \ldots, j_k) \in \{0, 1, \ldots, d\}^k, \ k = 0, 1, 2, \ldots \]

**Definition 2.1.** The set of all finite multi-indices is denoted by $A$. Given a multi-index $J = (j_1, \ldots, j_k)$ we define

(i) the length $|J|$ of the multi-index $J$ by $|J| = |(j_1, \ldots, j_k)| := k$.

(ii) the function $\| \cdot \| : A \to \mathbb{N}$ - referred to as the degree of a multi-index - as $\|J\| = \|(j_1, \ldots, j_k)\| := |J| + \text{ card}\{ j_i = 0 \mid 1 \leq i \leq k \}$

(iii) the concatenation of multi-indices by $\circ : A \times A \to A$:

\[ (j_1, \ldots, j_k) \circ (i_1, \ldots, i_l) = (j_1, \ldots, j_k, i_1, \ldots, i_l) \]

(iv) and finally the right and left decrement respectively as

\[ J^- = (j_1, \ldots, j_k)^- := (j_1, \ldots, j_{k-1}) \]

\[ -J = -(j_1, \ldots, j_k) := (j_2, \ldots, j_k) \]

2.2. Vector fields and other differential operators. In some cases, the functions $V_i$, $i = 0, \ldots, d$ are regarded as vector fields (or first order differential operators) on $\mathbb{R}^n$, i.e.

\[ V_i(f) = \sum_{j=1}^{n} V_i^j \frac{\partial}{\partial x_j} f \]

where $V_i^j$ denotes the $j$th coordinate function of $V_i$ and $f$ is a smooth $\mathbb{R}^n \to \mathbb{R}$ function. The identity function on $\mathbb{R}^n$ is denoted by $H$ and $V_i(H)$ is understood coordinate-wise.

By compositions of coefficient functions, we will refer to the operator composition, e.g. $V_i \circ V_j$ is defined

\[ V_i \circ V_k(f) = \sum_{j=1}^{n} V_i^j \frac{\partial}{\partial x_j} \left( \sum_{l=1}^{n} V_k^l \frac{\partial}{\partial x_l}(f) \right) \]

for a smooth function $f : \mathbb{R}^n \to \mathbb{R}$.

Given a multi-index $J = (j_1, \ldots, j_k)$, we introduce the notation

\[ V_J = V_{j_1} \circ \cdots \circ V_{j_k}. \]

2.3. Stratonovich and Itô iterated integrals. In this section we review some properties of the Stratonovich and the Itô integrals.

**Definition 2.2.** Given two continuous semi-martingales $X_t$ and $Y_t$, we define the Stratonovich integral as

\[ \int X \circ dY = \int X dY + \frac{1}{2} \langle X, Y \rangle \]

where $\int X dY$ denotes the Itô integral and $\langle X, Y \rangle$ is the cross-variation process.
To simplify the notation, we introduce $B^0_0 := t$, i.e. the 0th coordinate of the Brownian motion is the time. One collection of basic objects in the paper is given by the Stratonovich iterated integrals.

**Definition 2.3.** Given a multi-index $J = (j_1, \ldots, j_k)$ we define the Stratonovich iterated integrals as

$$B^J_{s,t} := \int_{s < t_1 < \cdots < t_k < t} \circ dB_{t_1}^{j_1} \circ \cdots \circ dB_{t_k}^{j_k}$$

$$B^J_{s,t}(Y) := \int_{s < t_1 < \cdots < t_k < t} Y_{t_1} \circ dB_{t_1}^{j_1} \circ \cdots \circ dB_{t_k}^{j_k}$$

for an integrable process $Y_t$.

The analogous Itô iterated integrals are denoted by $D^J_{s,t}$ and $D^J_{t,s}(Y)$ respectively.

We will mainly work with Stratonovich iterated integrals, but for some cases we will need the Itô forms. The Stratonovich iterated integrals considered in this paper can be represented by a linear combinations of Itô integrals. To derive the precise Stratonovich-Itô transformation we need the following relation between multi-indices.

**Definition 2.4.** Let’s regard a partition $(I_1, \ldots, I_k)$ of a multi-index $J$, i.e. $J = J_1 \circ \cdots \circ J_k$ for some $k$ such that $\|J_i\| \leq 2$ for $i = 1, \ldots, k$. Let’s suppose there exists a multi-index $I$ which can be partitioned into sub-indices $I_1, \ldots, I_k$ such that $I = I_1 \circ \cdots \circ I_k$ and for each $i = 1, \ldots, k$ either $J_i = I_i$ with both length 1 or $J_i = (l,l)$ and $I_i = (0)$ for some $l \in \{1, \ldots, d\}$. Then we will say that $I$ is related to $J$ through the partitions $(I_1, \ldots, I_k)$ and $(J_1, \ldots, J_k)$ of length $k$ and denote this relationship by

$$J \sim_k I$$

where $k$ denotes the number of subsets in the related partitions. Note that the relation $\sim_k$ is one to many.

If $J \sim_k I$, we define

$$\nu(J, I) := \text{card}\{i \mid 1 \leq i \leq k, \ J_i \neq I_i\}$$

where $(J_1, \ldots, J_k)$ and $(I_1, \ldots, I_k)$ denotes the partitions of $J$ and $I$ relating the two multi-indices.

Applying the definition of the Stratonovich integral, one can derive

$$\int_0^t \int_0^s \circ dB_u^i \circ dB_u^j = \int_0^t \int_0^s dB_u^i dB_u^j + \delta_{i,j} (1 - \delta_{i,0}) \frac{1}{2} \int_0^t dB_0$$

where $\delta_{i,j} = 1$ if $i = j$ and 0 otherwise. Recall, that for $0 \leq i, j \leq d$ and an integrable process $X$, the following equality holds.

$$\langle \int_0^t X_u dB_u^i, B^j_t \rangle_t = \delta_{i,j} (1 - \delta_{i,0}) \int_0^t X_u dB_u^0.$$  

Using the equalities (2.4) and (2.5) repeatedly, we get the following general formula.

**Lemma 2.1.** For any multi-index $J$

$$B^J_{s,t} = \sum_{J \sim_k I} \frac{1}{2^{\nu(J,I)}} D^I_{s,t}.$$
Let \( f \) be a smooth \( \mathbb{R} \rightarrow \mathbb{R} \) function. Recall, that Itô’s lemma (see equality (3.2)) implies:

\[
\langle f(\xi), B_t^i \rangle = \int_0^t V_i f(\xi_u) dB_u^0.
\]

Combining this equality with the rule (2.1), we get

\[
\int_0^t f(\xi_u) \circ dB^i_u = \int_0^t f(\xi_u) dB^i_u + (1 - \delta_{i,0}) \frac{1}{2} \int_0^t V_i f(\xi_u) dB^0_u.
\]

The lemma is implied:

**Lemma 2.2.** For any multi-index \( J \)

\[
B^I_{s,t}(f(\xi)) = \sum_{J \sim I} \frac{1}{2^{\nu(J,I)}} D^I_{s,t}(f(\xi)) \\
+ (1 - \delta_{j_1,0}) \frac{1}{2} \sum_{-J \sim I} \frac{1}{2^{\nu(J,I)}} D^I_{s,t}(D^J_{s,t}(V_{j_1} f(\xi))).
\]

**Proof.** First, we apply equation (2.7) to the inner-most term:

\[
B^I_{s,t}(f(\xi)) = B^I_{s,t} \left( \int_s^t f(\xi_u) dB^I_u \right) \\
+ (1 - \delta_{j_1,0}) \frac{1}{2} B^I_{s,t} \left( \int_s^t V_{j_1} f(\xi_u) dB^0_u \right).
\]

Then, iteratively we continue applying the rule (2.1) and in particular (2.5) to the inner-most Stratonovich terms in each iterated integral.

**Remark 2.1.** Note that the Itô form of a Stratonovich iterated integral of the type (2.2) or (2.3) corresponding to a multi-index \( J \) has a term corresponding to an all-zero multi-index if and only if \( J = J_1 \circ \cdots \circ J_k \) such that either \( J_i = (0) \) or \( J_i = (l, l) \) for all \( i = 1, \ldots, k \) and for some \( 1 \leq l \leq d \). In the case, when the Itô form of a Stratonovich iterated integral has no term corresponding to an all-zero multi-index, its expected value is zero.

The iterative use of Itô’s isometry and the Hölder inequality imply the \( L^2 \) boundedness of the Itô iterated integrals:

**Lemma 2.3.** There exist constants \( C_1 \) and \( C_2 \) depending on the multi-index \( J \) only, such that

\[
E \left[ |D^I_{s,t}(f(\xi))|^2 \right]^{1/2} \leq C_1 (t - s)^{\|J\|/2} \\
E \left[ |D^I_{s,t}(f(\xi))|^2 \right]^{1/2} \leq C_2 \|f\|_\infty (t - s)^{\|J\|/2}
\]

assuming that \( \|f\|_\infty \) is finite.

For the Stratonovich counterpart of the previous lemma we introduce the norm \( \|\cdot\|_\infty^V \) defined by

\[
\|f\|_\infty^V = \max \{ \|f\|_\infty, \|V_1 f\|_\infty, \ldots, \|V_d f\|_\infty \}.
\]
Lemma 2.4. There exist constants $C_1$ and $C_2$ depending on the multi-index $J$ only, such that

$$
E\left[|B^{J}_{s,t}|^2\right]^{1/2} \leq C_1(t-s)^{||J||/2}
$$

$$
E\left[|B^{J}_{s,t}(f(\xi))|^2\right]^{1/2} \leq C_2\|f\|_{\infty}^V \left[(t-s)^{||J||/2} + (t-s)^{(||J||+1)/2}\right]
$$

assuming that $\|f\|_{\infty}^V$ is finite.

The error of the numerical schemes we consider in this paper can be represented in terms of iterated integrals. The following lemma is the key step for determining an upper bound for the global error.

Lemma 2.5. For each multi-index $J$, there exists a constant $C_J$ not depending on $k$ such that for any Borel function $f$ satisfying $\sup_{t \in [0,T]} E \left[\|f(\xi)\|^2\right] < \infty$, the following inequality holds

$$
E\left[\sup_{1 \leq l \leq k} \left(\sum_{i=1}^{l} D_{t_{i-1},t_i}(f(\xi))\right)^2\right]^{1/2} \leq C_J \Delta t^m \|f\|_{\infty}
$$

where

(i) $m = (||J|| - 2)/2$ if $||J|| = |J|$ or $J = (j,j)$ for some $j$, $1 \leq j \leq d$

(ii) $m = (||J|| - 1)/2$ otherwise

and furthermore $\Delta t > 0$ denotes the longest step length in the partition $0 = t_0 < t_1 < \cdots < t_k \leq T$.

One can derive the proof for the above Lemma by applying Lemma 2.3 and Doob’s inequality, exploiting the fact that in case $||J|| \neq |J|$ the terms under summation (2.8) are orthogonal. A detailed proof can be found in Chapter 10 of [9].

The Stratonovich version of Lemma 2.5 is as follows.

Lemma 2.6. For each multi-index $J$ there exists a constant $C_J$ not depending on $k$ such that for any smooth function $f$ as in Lemma 2.5, the following inequality holds

$$
E\left[\sup_{1 \leq l \leq k} \left(\sum_{i=1}^{l} B^{J}_{t_{i-1},t_i}(f(\xi))\right)^2\right]^{1/2} < C_J \left\{K_J \Delta t^{m-1/2} + \Delta t^m\right\} \|f\|_{\infty}^V
$$

where $m = (||J|| - 1)/2$ and

(i) $K_J = 1$ if $J$ is of the form $J_1 \circ \cdots \circ J_s$ for some $s$ such that for all $i = 1, \ldots, s$ either $J_i = (0)$ or $J_i = (j,j)$ for some $j$, $1 \leq j \leq d$

(ii) $K_J = 0$ otherwise

furthermore $\Delta t \in (0,1]$ denotes the length of the longest step in the partition $0 = t_0 < t_1 < \cdots < t_k \leq T$.

Proof. There exist constants $C_1$ and $C_2$ such that using Lemma 2.2

$$
E\left[\left(\sup_{1 \leq l \leq k} \sum_{i=1}^{l} B^{J}_{t_{i-1},t_i}(f(\xi))\right)^2\right]
$$
\[ \leq C_1 \sum_{J \in \mathcal{I}_0} \mathbb{E} \left[ \sup_{1 \leq l \leq k} \left( \sum_{i=1}^{l} \frac{1}{2^{\nu(J,I)}} D_{t_{i-1},t_i}^l (f(\xi)) \right)^2 \right] \]

\[ + C_2 \sum_{J \in \mathcal{I}_0} \mathbb{E} \left[ \sup_{1 \leq l \leq k} \left( \sum_{i=1}^{l} \frac{1}{2^{\nu(J,I)}} D_{t_{i-1},t_i}^l (D^0_{t_{i-1},j} f(\xi)) \right)^2 \right] \]

Then Lemma 2.5 and Remark 2.1 imply the assertion. \[
\square
\]

**Remark 2.2.** Note that Remark (2.1) implies that if in the previous lemma the degree of the multi-index \( J \) is odd then \( K_J = 0 \) in (2.10) and the bound is of order \( m = (\|J\| - 1)/2 \).

The following lemma is a special case of Lemma 2.6.

**Lemma 2.7.** For each multi-index \( J \) there exists a constant \( C_J \) not depending on \( k \) such that for any Borel function \( f \) as in Lemma 2.5, the following inequality holds

\[
(2.10) \quad \mathbb{E} \left[ \left( \sup_{1 \leq l \leq k} \sum_{i=1}^{l} f(\xi_{t_{i-1}}) B_{t_{i-1},t_i}^l \right)^{2} \right] ^{1/2} < C_J \left\{ K_J \Delta t^m -1/2 + \Delta t^m \right\} \|f\|_{\infty}
\]

where \( m = (\|J\| - 1)/2 \) and

(i) \( K_J = 1 \) if \( J \) is of the form \( J_1 \circ \cdots \circ J_s \) for some \( s \) such that for all \( i = 1, \ldots, s \) either \( J_i = (0) \) or \( J_i = (j,j) \) for some \( j, 1 \leq j \leq d \) and

(ii) \( K_J = 0 \) otherwise

furthermore \( \Delta t \in (0,1] \) denotes the length of the longest step in the partition \( 0 = t_0 < t_1 < \cdots < t_k \leq T \).

**Lemma 2.8.** Given the multi-indices \( J_1, \ldots, J_k \) there exists a constant not depending on \( s \) and \( t \) such that

\[
\mathbb{E} \left[ \prod_{i=1}^{k} B_{s,t}^{J_i} \right] ^{1/2} < C(t-s)^{(1/2) \sum_{i=1}^{k} \|J_i\|}
\]

The lemma is a corollary of Lemma 2.2, Lemma 2.3 and the fact that for the multi-indices \( I = (i_1, \ldots, i_q) \) and \( J = (j_1, \ldots, j_r) \)

\[
B_{s,t}^{I} B_{s,t}^{J} = \int_{s\leq u_1<\cdots<u_q<s} \cdots \int_{s\leq v_1<\cdots<v_r<s} \od B_{u_1}^{i_1} \cdots \od B_{u_q}^{i_q} \od B_{v_1}^{j_1} \cdots \od B_{v_r}^{j_r}
\]

\[
= \sum_{\sigma \in \text{Shuffles}(q,r)} \int_{s\leq u_1<\cdots<u_q<s} \cdots \int_{s\leq v_1<\cdots<v_r<s} \od B_{u_1}^{\sigma^{-1}(1)} \cdots \od B_{u_q}^{\sigma^{-1}(q+r)}
\]

where \( \text{Shuffles}(q,r) \) is the set of all permutations of the set \( \{1,2,\ldots,q+r\} \) such that \( \sigma \in \text{Shuffles}(q,r) \) if and only if

\[
\sigma(1) < \sigma(2) < \cdots < \sigma(q) \text{ and } \sigma(q+1) < \sigma(q+2) < \cdots < \sigma(q+r).
\]

For more details on shuffles, we refer the reader to [17].
3. The log-signature theorem

In the log-signature theorem, a random ordinary differential equation is derived which approximates the solution to the SDE locally. In some forms, this theorem was proved in [1], [4] and [8] and in addition, a local error formula was given. The log-signature theorem is also an extension of the result in [18] to non-linear and stochastic differential equations. In this section the theorem is reformulated in terms of the log-signature of the driving noise and a more specific local error formula is also derived.

3.1. Stochastic Taylor expansion. The expansion presented in this section is a stochastic extension of the deterministic Taylor expansion using Itô’s differentiation rule. Given the SDE

\[ \begin{align*} 
\frac{d\xi_t}{\xi_0} &= \sum_{i=0}^{d} V_i(\xi_t) \circ dB_i^t \\
\end{align*} \]

the Itô lemma in Stratonovich form is

\[ f(\xi_t) = f(\xi_0) + \sum_{i=0}^{d} \int_0^t V_i(f(\xi_s)) \circ dB_i^s. \]

Assuming sufficient smoothness of \( f \), applying Itô’s lemma to the terms under the integral leads to the equality

\[ f(\xi_t) = f(\xi_0) + \sum_{i=1}^{d} \int_0^t V_i(f(\xi_s)) \circ dB_i^s + \]

\[ + \int_0^t V_0(f(\xi_s)) \circ dB_0^s + \sum_{i,j=0}^{d} \int_0^t \int_0^s V_j(V_i(f(\xi_u))) \circ dB_i^u \circ dB_j^s. \]

By the iterative use of Itô’s lemma, one can derive the stochastic Taylor expansion of \( f(\xi_t) \) as follows.

\[ f(\xi_t) = f(\xi_0) + \sum_{\|J\| \leq m} V_J(f(\xi_0)) B_{0,t}^J + \sum_{\|J\| = m} B_{0,t}^J(V_J(f(\xi))). \]

The key rule in the \( m \)th step of the expansion is to expand only those terms which correspond to multi-indices of degree \( m + 1 \).

Similarly, one can derive the stochastic Taylor expansion of \( \xi_t \) by applying (3.3) coordinate-wise on the identity function \( H \) of \( \mathbb{R}^n \):

**Lemma 3.1.**

\[ \xi_t = \xi_0 + \sum_{\|J\| \leq m} V_J(\xi_0) B_{0,t}^J + \sum_{\|J\| = m} B_{0,t}^J(V_J(\xi)). \]
where the last term is referred to as the remainder and denoted by $R_{m_{0},t}^m$. Furthermore, there exists a constant $C$ not depending on $t$ such that

$$
\mathbb{E} \left[ \left\| R_{m_{0},t}^m \right\|^2 \right]^{1/2} \leq C \sum_{\|J\|=m} t^{(m+1)/2} \left\| V_{J} H \right\|_{\infty}^{V} 
+ C \sum_{\|J\|=m} t^{(m+2)/2} \left\| V_{J} H \right\|_{\infty}^{V}
$$

**Proof.** The $L^2$ bound on the remainder term is implied by Lemma 2.2 and Lemma 2.3. \(\square\)

### 3.2. Algebraic setting.

Lemma 3.1 demonstrates that the stochastic Taylor expansion without remainder term gives a good approximation of the solution to the corresponding SDE on short time steps. It also demonstrates the importance of the information provided by the set of iterated integrals on the time interval $[0,t]$ corresponding to multi-indices of norm at most $m$. As it was pointed out in [14] and [15], these iterated integrals can be formulated into an algebraic object called the Brownian signature. The signature is a random element in the tensor algebra $\mathcal{T}$ defined as

**Definition 3.1.** $\mathcal{T}$ denotes the non-commutative tensor algebra generated by the letters $\varepsilon_0, \varepsilon_1, \ldots, \varepsilon_d$. For a multi-index $J = (j_1, \ldots, j_k)$ we introduce the notation

$$
\varepsilon_J = \varepsilon_{j_1} \otimes \cdots \otimes \varepsilon_{j_k}
$$

Furthermore, the empty word is denoted by $1$.

A general element in $\mathcal{T}$ is of the form

$$
a = \sum_{J \in A} a_J \varepsilon_J
$$

with real coefficients $a_J$, $J \in A$. We define the operations in $\mathcal{T}$ as follows.

**Definition 3.2.** The multiplication of the tensor elements is written as

$$
a \otimes b = \sum_{J, I \in A} a_J b_I \varepsilon_{J \circ I}
$$

$$
a^\otimes k := a \otimes \cdots \otimes a \quad \text{k times}
$$

The exponential of a tensor element is defined by the power series

$$
\exp(a) := 1 + \sum_{k \geq 1} \frac{1}{k!} a^\otimes k
$$

If $a = 1 + \sum_{J \in A, |J| > 0} a_J \varepsilon_J$, then we define the logarithm of $a$ as

$$
\log(a) := \sum_{k \geq 1} \frac{(1-a)^k}{k} (-1)^k
$$

We denote the simplex $\{(s, t) \mid 0 \leq s \leq t \leq T\}$ by $\Delta_T$. 
Definition 3.3. (Signature, log-signature) The Stratonovich Brownian signature is a $\Delta_T \to T$ function, defined as

$$S_{s,t}(\circ B) := 1 + \sum_{J \in A} B^J_{s,t} \varepsilon_J$$

the log-signature is defined as

$$L_{s,t} := \log(S_{s,t}).$$

Definition 3.4. We define the Lie bracket of the elements $a$ and $b$ as $[a, b] := a \otimes b - b \otimes a$. The vector space spanned by all finite lie brackets formed by the letters $\varepsilon_0, \ldots, \varepsilon_d$ is the Lie algebra $\mathcal{L}$. The elements of the closure $\overline{\mathcal{L}}$ are referred to as Lie series.

For further algebraic background we refer the reader to [21].

The proof for following properties of the Brownian signature can be found in [15].

Lemma 3.2. (Properties of the Brownian signature)

(i) The signature is multiplicative, i.e. for $0 \leq s \leq t \leq u \leq T$

$$S_{s,t} \otimes S_{t,u} = S_{s,u}$$

(ii) The log signature $L_{s,t}$ is a.s. a Lie series.

(iii) $\exp(L_{s,t}) = S_{s,t}$

Lemma 3.2 has crucial importance in constructing numerical methods and in deriving simpler proof for the main result of [4]. In particular, the fact that the (truncated) logarithm of the Brownian signature is a.s. a Lie series (Lie algebra element) and the fact that the exponential of the log-signature is the signature will play a central role when proving the Log-signature theorem (Theorem 3.1).

Definition 3.5. We also introduce the truncation operator $\pi_m$ defined as

$$\pi_m \sum_{J \in A} a_J \varepsilon_J := \sum_{J \in A \; \|J\| \leq m} a_J \varepsilon_J.$$

Definition 3.6. Finally we introduce an algebra homomorphism $\Gamma$ mapping tensor algebra elements to differential operators generated by $\Gamma(\varepsilon_i) := V_i, \; i = 0, 1, \ldots, d$.

Note that the function $\Gamma$ maps Lie algebra elements to vector fields. For example $\Gamma(\varepsilon_i) = V_i, \; \Gamma(\varepsilon_0 + [\varepsilon_i, \varepsilon_j]) = V_0 + [V_i, V_j]$, etc.

3.3. Proof of the log-signature theorem. Given the algebraic setting of the previous section we rephrase the result of [1], [4] and [8] in terms of the log-signature. We apply some ideas from [10] and [11].

Given a vector field $W$, the solution for the ordinary differential equation

$$\begin{align*}
\frac{dy_t}{dt} &= W(y_t)dt \\
y_0 &= x_0
\end{align*}$$

is denoted by

$$y_t = \text{Exp}(tW)(y_0).$$
Theorem 3.1. (Log-signature theorem) Given the SDE (3.1) with smooth coefficient functions \(V_0, \ldots, V_d\) with bounded derivatives up to order \(2m+1\), and a \(t \in (0,1)\), there exists a constant \(C_1\) not depending on \(t\) but on \(m\), such that

\[
E \left[ \| R_1 \|_2 \right]^{1/2} \leq C_1 \sum_{m+1 \leq \| I \| \leq 2m} t^{\| I \|/2} \| V_I H \|_\infty^V
\]

where

\[
R_1 = \xi_t - y_1, \quad y_1 = \text{Exp} \left( \Gamma(\pi_m L_{0,t}) \right)(x_0), \quad L_{0,t} = \log[S_{0,t}(\circ B)].
\]

Furthermore there exists a constant \(C_2\) not depending on \(t\) but on \(m\), satisfying

\[
E \left[ \| R_2 \|_2 \right]^{1/2} \leq C_2 \sum_{m+2 \leq \| I \| \leq 2m} t^{\| I \|/2} \| V_I H \|_\infty^V
\]

where

\[
R_2 = R_1 - \Gamma \left[ (\pi_{m+1} - \pi_m) L_{0,t} \right] (\xi_0).
\]

Remark 3.1. Note that the differential operator defined as \(W := \Gamma(\pi_m L_{0,t})\) is a.s. the projection of a Lie algebra element and hence a vector field.

Proof. Applying the notation introduced in the previous section, the stochastic Taylor expansion can also be represented as the image of the signature under the algebra homomorphism

\[
\xi_t = \Gamma \left[ \pi_m S_{0,t}(\circ B) \right] (\xi_0) + R^m_{0,t}.
\]

Since \(\exp(L_{0,t}) = S_{0,t}(\circ B)\), the stochastic Taylor expansion (3.4) can be rewritten as

\[
\xi_t = \Gamma \left[ \pi_m \sum_{j=0}^{m} \frac{1}{j!} \left( L_{0,t} \right)^\otimes j \right] (\xi_0) + R^m_{0,t}.
\]

On the other hand, we derive a deterministic Taylor-type expansion of \(y_1\) as follows. For any differentiable function \(f : \mathbb{R}^n \to \mathbb{R}\)

\[
df(y_t) = W(f(y_t))dt.
\]

Applying (3.11) coordinate-wise on the identity function \(H\) of \(\mathbb{R}^n\), we get

\[
y_1 = y_0 + \int_0^1 W \circ H(y_t) dt
\]

(3.12) \[= y_0 + W(y_0) + \int_{0<s_1<s_2<1} W \circ H(y_{s_1}) ds_1 ds_2.\]

By Definition 3.2 of the log function, the vector field \(W\) can be represented as a finite linear combination of the functions \(V_J\) of the form

\[
W = \sum_{\| J \| \leq m} \left( \sum_{\| I \| \leq |J|} a_{J_1,\ldots,J_l} B^{J_1}_{0,t} \cdots B^{J_l}_{0,t} \right) V_J
\]
where the real coefficients $a_{J_1,...,J_l}$ only depend on the multi-indices $J_1,\ldots,J_l$. We introduce the notation

$$A^J_{s,t} = \sum_{J_1 \circ \cdots \circ J_l = J} a_{J_1,...,J_l} B_{s,t}^{J_1} \cdots B_{s,t}^{J_l}.$$ 

The expression under the integral in (3.12) can be written as a weighted sum of functions where the weights are products of $A^J_{0,t}$-type terms. If the weight of a term is $A^J_{1,0} \cdots A^J_{k,0}$ for some multi-indices $J_1,\ldots,J_k$, we will say that the degree of the term is $\|J_1\| + \cdots + \|J_k\|$. Applying (3.11) repeatedly on the terms under the integral sign with degree at most $m$, one can derive the following formula

$$y_1 = \Gamma \left[ \pi_m \sum_{j=0}^{m} \frac{1}{j!} (\pi_m L_{0,t})^o^j \right] (y_0) + \tilde{R}_{0,t}^m.$$ 

Lemma 2.8 implies

$$E \left[ \left\| \tilde{R}_{0,t}^m \right\|^2 \right]^{1/2} < C \sum_{\text{all } ||I|| \leq 2m} t^{||I||/2} \left\| V_{I} H \right\|_{\infty}.$$ 

for some constant not depending on $t$.

Comparing the solution to the SDE (3.10) and the solution to the ODE (3.13), the difference is given by $\xi_t - y_1 = R_{0,t}^m - \tilde{R}_{0,t}^m$. Given the stochastic and deterministic Taylor expansions, $R_{0,t}^m$ contains terms of degree $m + 1$ and $m + 2$. Hence the inequality (3.14) and Lemma 2.3 imply (3.6), (3.8) and (3.7).

Theorem 3.1 shows the importance of locally derived random ODEs.

**Definition 3.7.** We call a numerical approximation of the solution to the SDE (3.1) **ODE-approach based approximation corresponding to the $m$-truncated log-signature** if at each step a random ODE is derived from $\pi_m L_{s,t}$ and solved or numerically approximated, i.e. for $0 = t_0 < t_1 < \cdots < t_k = T$ and $\hat{X}_{t_0} = \xi_0$

$$\hat{X}_{t_{i+1}} = \text{Exp}(\Gamma(\pi_m L_{t_i,t_{i+1}}))(\hat{X}_{t_i}),$$

where $\text{Exp}(W)(x)$ denotes a numerical approximation to $\text{Exp}(W)(x)$ for any vector field $W$ and $x \in \mathbb{R}^n$, provided the solution exists.

Note that to achieve high order global convergence, the use of a sufficiently accurate ODE numerical method is required. We specify the required accuracy in section 4. For numerical methods approximating the solution to ODEs, we refer the reader to the monograph [2] by Butcher.

**Remark 3.2.** One of the most important advantages of the ODE approach based schemes implemented with highly accurate ODE solvers over the purely stochastic Taylor expansion based schemes is the stability in the following sense. The ODE approach approximates $\xi_t$ with points which are reachable by the diffusion. E.g. if the diffusion a.s. does not leave a subset of $\mathbb{R}^n$, the ODE approach results in points from or very close to this subset. The purely stochastic Taylor expansion based schemes do not have this property.
3.4. The nilpotent case. In this section we assume that there exists a positive integer $m$ such that the coefficient functions $V_i$, $i = 0, \ldots, d$ satisfy the $m$-nilpotent condition, i.e.

**Definition 3.8.** Let’s suppose that there exist a positive integer $m$ such that any bracket of the letters $\varepsilon_{i_1}, \ldots, \varepsilon_{i_k}$ (e.g. $[\varepsilon_{i_1}, \ldots, [\varepsilon_{i_{k-1}}, \varepsilon_{i_k}], \ldots]$ etc.), $i_j \in \{0, \ldots, d\}$ for $j = 1, \ldots, k$ satisfying $\| (i_1, \ldots, i_k) \| \geq m$ is projected to the zero operator by the function $\Gamma$. Then we say that the coefficient functions $V_i$, $i = 0, \ldots, d$ satisfy the $m$-nilpotent condition.

In this case, $\Gamma[\pi_k L_{0,t}] = \Gamma[L_{0,t}]$ for all $k \geq m$, in particular

(3.15) \quad \Gamma[\pi_k S_{0,t}(\circ B)] = \Gamma[\pi_k \exp(L_{0,t})] = \Gamma[\pi_k \exp(\pi_m L_{0,t})] = \Gamma \left[ \pi_k \sum_{j=0}^{\infty} \frac{1}{j!} \pi_m L_{0,t}^j \right]

Hence, we have the following.

**Lemma 3.3.** Let’s assume that the coefficient functions of the SDE (3.1) satisfy the $m$-nilpotent condition for some positive integer $m$. If for a given $x_0$ the series

(3.16) \quad \Gamma[S_{0,t}(\circ B)](x_0) = \Gamma \left[ \sum_{j=0}^{\infty} \frac{1}{j!} \pi_m L_{0,t}^j \right] (x_0) = \sum_{j=0}^{\infty} \frac{1}{j!} \Gamma[\pi_m L_{0,t}]^j (x_0)

is convergent, then

(3.17) \quad \xi_t = \exp(\Gamma(\pi_m L_{0,t}))(x_0) = \exp(\Gamma(L_{0,t}))(x_0)

is satisfied without any remainder term.

**Example 3.1.** From the practical point of view, Lemma 3.3 tells us that given the vector field corresponding to the $m$-truncated log-signature, the solution for the derived ODE results in an exact solution for the SDE at time $t$. Note that it is only exact at time $t$ but not inside the interval $[0, t]$.

One simple example for the nilpotent case is given by the geometric Brownian motion (in Stratonovich form)

\[ d\xi_t = \left( \mu - \frac{1}{2} \sigma^2 \right) \xi_t dt + \sigma \xi_t \circ dB \]

which is 2-nilpotent. Lemma 3.3 states that

(3.18) \quad \xi_t = \exp[\Gamma(\varepsilon_0 t + \varepsilon_1 B_{0,t})](\xi_0) = \exp(W)(\xi_0).

where \[ W(x) = ((\mu - \sigma^2/2)t + \sigma B_{0,t})x \]

On the other hand, it is well known that \[ \xi_t = \xi_0 e^{(\mu - \frac{1}{2} \sigma^2)t + \sigma B_{0,t}} \]

i.e. the Brownian increment determines the solution, which is equivalent to (3.18).
4. ODE-approach based schemes

In this section, we focus on the global error of discretization schemes based on the ODE-approach given by a partition \( 0 = t_0 < \cdots < t_k = T \) and a random sequence \( \xi_0 = \hat{X}_{t_0}, \hat{X}_{t_1}, \ldots, \hat{X}_{t_k} \) approximating the solution to the SDE (3.1) (see Definition 3.7). First we consider the case, when the exact solution of the random ODEs are used. Then we analyse the case, when the solutions to the random ODEs are numerically approximated. We derive the required accuracy of the numerical solvers to preserve the global order of convergence achieved in the corresponding version of the former case. Finally, we highlight how some of the well known schemes fit into the ODE-approach based framework.

4.1. From local to global error. We define the local error of the scheme as follows.

**Definition 4.1.** Given a discretization scheme by a random sequence \( \hat{X}_{t_0}, \ldots, \hat{X}_{t_k} \) approximating the solution to the SDE (3.1), for \( i = 1, \ldots, k \) we define the local error corresponding to the time interval \([t_{i-1}, t_i]\) as \( E_i := \hat{X}_{t_i} - Y_i \) where \( Y_i \) is the solution to the SDE

\[
\begin{align*}
    dY^i_t &= \sum_{i=0}^d V_i(Y^i_t) \circ dB^i_t \\
    Y^i_{t_{i-1}} &= \hat{X}_{t_{i-1}}
\end{align*}
\]

The global error of the same scheme corresponding to the time interval \([0, t_k]\) is defined as \( \xi_{t_k} - \hat{X}_{t_k} \).

The local error of discretization schemes based on the exact solution of the ODE derived from the log-signature of the driving noise is given by Theorem 3.1. The practical discretization schemes are derived by choosing a (numerical) solver for the ODE. However, from Theorem 3.1 the order of the global convergence is not obvious and also the accuracy required from the numerical ODE solver to ensure the global convergence is unclear.

Firstly, we give a sufficient condition on the local error for the global convergence using similar techniques as in [5] and [9]. Then we show when a numerical ODE solver satisfies this condition. The proof of Theorem 4.1 applies the following version of Gronwall’s Lemma.

**Lemma 4.1. (Gronwall’s Lemma)** Let \( s, t \in \mathbb{R}, s < t \) and suppose that \( b, c \) and \( r \) are \( \mathbb{R} \to \mathbb{R} \) functions such that \( b \) is continuous, \( c \) is continuously differentiable and \( r \) is piece-wise continuous on \([s, t]\), furthermore

\[
r(u) \leq c(u) + \int_s^u b(v)r(v)dv
\]

Then

\[
r(u) \leq c(u) + \int_s^u b(v)c(v)e^{\int_s^u b}dv
\]

**Theorem 4.1. (Bound on the global error)** Let’s suppose that a discretization scheme \( \hat{X}_{t_0}, \ldots, \hat{X}_{t_k} \) approximating the solution to the SDE (3.1) and corresponding to the partition \( 0 = t_0 < t_1 < \cdots < t_k = T \) with longest time step
\( \Delta t \in (0, 1) \) has local error \( E_i \) \( i = 1, \ldots, k \) as in Definition 4.1, such that

\[
E \left[ \sup_{1 \leq l \leq k} \left\| \sum_{i=1}^{l} E_i \right\|^2 \right]^{1/2} < C_1 \Delta t^{m/2}
\]

for some positive constant \( C_1 \) not depending on \( \Delta t \). Furthermore suppose that the coefficient function \( V_0 \) is continuously differentiable with bounded gradient and the functions \( V_1, \ldots, V_d \) are twice continuously differentiable with bounded first and second derivatives. Let \( K_V \) be defined by

\[
K_V = \max \left\{ \|
abla V_i H\|_\infty, i = 0, \ldots, d \right\} \cup \left\{ \|
abla V_j V_i H\|_\infty, i, j = 1, \ldots, d \right\}.
\]

Then there exists a constant \( C_2 \) depending on \( K_V, C_1 \) and \( T \) but not depending on \( \Delta t \) such that

\[
E \left[ \| \xi_T - \tilde{X}_T \|_2 \right]^{1/2} < C_2 \Delta t^{m/2}.
\]

**Proof.** Given the partition \( 0 = t_0 < t_1 < \cdots < t_k = T \), for each \( 1 \leq i \leq k \) we consider the process \( Y_{i,t} \) introduced in Definition 4.1. Note that \( Y_{i,t} \) can be equivalently represented as

\[
Y_{i,t}^i = \tilde{X}_{t_{i-1}} + \sum_{j=0}^{d} \int_{t_{i-1}}^{t_i} V_j(Y_{i,s}) \circ dB_j^i
\]

\[
Y_{i,t}^i = \tilde{X}_{t_i} - E_i.
\]

We also define

\[
Y_t := Y_t^i \quad \text{for } t \in [t_{i-1}, t_i), \quad i = 1, \ldots, k,
\]

see Figure 1 for visualisation. Note that for \( t \in [0, T) \), \( Y \) satisfies

\[
Y_t = \sum_{j=1}^{d} \int_{0}^{t} V_j(Y_s) \circ dB_j^s + \sum_{i|t_i < t} E_i.
\]

Equation (4.3) implies

\[
E \left[ \| \xi_T - Y_t \|_2 \right]^{1/2} \leq E \left[ \left\| \sum_{i=0}^{d} \int_{0}^{t} V_i(\xi_s) - V_i(Y_s) \circ dB_i^s \right\|^2 \right]^{1/2} + E \left[ \sup_{l|l \geq t_{i-1}} \left\| \sum_{j=1}^{l} E_j \right\|^2 \right]^{1/2}
\]

Let’s introduce the notation

\[
r(t) := E \left[ \| \xi_T - Y_t \|_2 \right]^{1/2}
\]

By Lemma 2.2, Itô’s isometry and the boundedness assumption on the coefficient functions and their derivatives, there exists a constant \( K_1 \) depending only on \( d \) such that

\[
E \left[ \left\| \sum_{i=0}^{d} \int_{0}^{t} V_i(\xi_s) - V_i(Y_s) \circ dB_i^s \right\|^2 \right]^{1/2} \leq K_1 K_V \int_{0}^{t} r(u) du.
\]
Then applying Gronwall’s lemma on the functions $r, b \equiv K_1 K_V$ and

$$c(t) := \mathbb{E} \left[ \sup_{t \geq t_{u-1}} \left( \sum_{j=1}^{t} E_j \right)^2 \right]^{1/2}$$

we get

$$\mathbb{E} \left[ \|\xi_T - Y_T\|^2 \right]^{1/2} \leq c(T) e^{TK_1 K_V}$$

Using the inequality (4.2) and defining $C_2 := C_1 e^{TK_1 K_V}$ the assertion is proved. □

**Corollary 4.1.** Let the sequence of local errors $E_i, i = 1, \ldots, k$ corresponding to the discretisation scheme $\tilde{X}_{t_0}, \ldots, \tilde{X}_{t_k}$ be given as in Theorem 4.1. Let’s suppose that for $i = 1, \ldots, k$, $E_i = M_i + N_i$ satisfying

(i) $\mathbb{E} \left[ \|M_i\|^2 \right]^{1/2} \leq C_1 \Delta t^{(m+1)/2}$ and $\mathbb{E} [M_i] = 0$

(ii) $\mathbb{E} \left[ \|N_i\|^2 \right]^{1/2} \leq C_1 \Delta t^{(m+2)/2}$

for some constant $C_1$ and $\Delta t \in (0, 1)$. Furthermore, assume that the coefficient functions satisfy the conditions of Theorem 4.1.

Then there exists a constant $C_2$ not depending on $\Delta t$ but depending on $C_1$ and $K_V$, such that

$$\mathbb{E} \left[ \|\xi_T - \tilde{X}_T\|^2 \right]^{1/2} < C_2 \Delta t^{m/2}$$
Proof. We only need to prove that \( \{ E_i \mid i = 1, \ldots, k \} \) satisfy the condition (4.2) of Theorem 4.1. Note that

\[
\left\| \sum_{i=1}^l E_i \right\|^2 \leq 2 \left\| \sum_{i=1}^l M_i \right\|^2 + 2 \left\| \sum_{i=1}^l N_i \right\|^2
\]

Given the assumptions on \( M_i \), the first term on the right hand side of (4.5) is a discrete time sub-martingale indexed by \( l \). Then using the conditions of the Corollary and applying Doob’s inequality the assertion is proved. \( \square \)

Corollary 4.2. Let’s assume that the coefficient functions \( V_0, \ldots, V_d \) satisfy the conditions of Theorem 3.1 for a positive even integer \( m \geq 2 \) and \( K_V \) is finite. Let’s suppose that a random sequence \( \hat{X}_{t_0}, \ldots, \hat{X}_{t_k} \) with longest step size \( \Delta t \in (0, 1) \) is the result of an ODE-approach based scheme derived from the \( m \)-truncated log-signature implemented with a numerical ODE solver satisfying

\[
E \| R_i \|^2 \leq C_1 \Delta t^{(m+2)/2}
\]

for some constant \( C_1 \).

Then there exists a constant \( C_2 \) not depending on \( \Delta t \), such that

\[
E \left\| \xi_T - \hat{X}_T \right\|^{1/2} < C_2 \Delta t^{m/2}.
\]

Proof. Since \( m \) is even, Remark 2.2 and Lemma 2.6 imply that defining

\[
M_i := \Gamma \left[ (\pi_{m+1} - \pi_m) L_{t_i, t_{i+1}} \right] (\xi_0) + R_i
\]

\[
N_i = R_i,
\]

the conditions of Corollary 4.1 are satisfied and so the assertion is proved. \( \square \)

Example 4.1. For any even integer \( m \) the discretization schemes defined by the \( m \)-truncated stochastic Taylor expansion satisfy the conditions of Corollary 4.2 and so they have a global order \( m/2 \).

4.2. Some examples and remarks. In general, the ODE-approach is based on random ODEs driven by the vector field \( W_{s,t}^m := \Gamma [\pi_m L_{s,t}] \) for some positive \( m \). Recalling Remark 2.2, the Itô form of the Stratonovich iterated integrals corresponding to the multi-index \( J \) has an all-zero multi-index corresponding integral if \( J = J_1 \circ \cdots \circ J_k \) such that either \( J_i = (0) \) or \( J_i = (l_i, l_i) \) for all \( i = 1, \ldots, k \) and for some \( 1 \leq l_i \leq d \). This implies that for an odd integer \( m \), there are integrals corresponding to all-zero multi-index in the Itô form of \( \Gamma [(\pi_{m+1} - \pi_m) L_{t_i, t_{i+1}}] (\xi_t) \) and hence the conditions of Corollary 4.2 are not necessarily satisfied, i.e. the global order is lower than \( m/2 \). In particular, when \( m = 1 \), some fixing is required to achieve convergence; we introduce an extended scheme as follows.

Definition 4.2. Given the time interval \([s, t]\) we introduce the random vector field

\[
\hat{W}_{s,t}^j := \Gamma [B_{s,t}^0 \xi_0 + \pi_1 L_{s,t}]
\]

which will be referred to as the vector field corresponding to the extended ODE approach based 1-truncated scheme.
Remark 4.1. Note that the Itô form of $\Gamma \left[ \pi_2 L_{s,t} - B^0_{s,t} \varepsilon_0 - \pi_1 L_{s,t} \right] (\xi_s)$ only includes stochastic integrals and so the conditions of Corollary 4.2 are satisfied. Therefore the global order of the extended ODE approach based 1-truncated scheme is in general 1 when $d = 1$ and $1/2$ for $d > 1$. In special cases, the global order might be higher: see Example 3.1 on the nilpotent case.

The schemes - probably the most commonly used in practice - are derived from $\hat{W}^1$ or $W^2$. One can regard both the Milstein scheme and the Heun scheme (ref.: [6], [9]) as particular versions of the ODE approach based schemes. In particular:

(i) the Heun scheme is equivalent to applying the predictor-corrector numerical scheme on the ODE driven by the vector field corresponding to the extended ODE approach based 1-truncated scheme:

$$
\hat{X}^\text{pred}_{t+1} = \hat{X}_t + \hat{W}^1_{t,t+1}(\hat{X}_t),
$$
$$
\hat{X}_{t+1} = \hat{X}_t + \frac{1}{2} \left[ \hat{W}^1_{t,t+1}(\hat{X}_t) + \hat{W}^1_{t+1,t+1}(\hat{X}^\text{pred}_{t+1}) \right].
$$

In general, the Heun scheme has global order 1/2.

(ii) for $d = 1$, the Milstein scheme is the truncated version of the Taylor-expansion based numerical approximation of the ODE driven by the vector field corresponding to the extended ODE approach based 1-truncated scheme:

$$
\begin{align*}
\hat{X}_{t+1} &= \hat{X}_t + \hat{W}^1_{t,t+1}(\hat{X}_t) + \frac{1}{2} \hat{W}^1_{t,t+1} \circ \hat{W}^1_{t,t+1} (\hat{X}_t) \\
&= V_0(\hat{X}_t) B^0_{t,t+1} + V_1(\hat{X}_t) B^1_{t,t+1} + \frac{1}{2} V_1 \circ V_1(\hat{X}_t) \left( B^1_{t,t+1} \right)^2 \\
&\quad + \frac{1}{2} \left( V_0 \circ V_1(\hat{X}_t) + V_1 \circ V_0(\hat{X}_t) \right) B^0_{t,t+1} B^1_{t,t+1}.
\end{align*}
$$

where the last line can be omitted from the actual Milstein scheme without losing the order of accuracy.

(iii) the Milstein scheme in higher dimensions is the truncated version of the two-step Taylor-expansion based numerical approximation of the ODE driven by the vector field

$$
W^2_{s,t} = \Gamma \left[ \pi_2 L_{s,t} \right] = \sum_{i=0}^{d} V_i B^i_{s,t} + \frac{1}{2} \sum_{2 \leq i < j \leq d} [V_i, V_j] \left( B^{(i,j)}_{s,t} - B^{(j,i)}_{s,t} \right).
$$

In general, the Milstein scheme has global order 1.

Remark 4.2. Note that by Corollary 4.2, the convergence of the extended ODE approach based 1-truncated scheme is due to the fact that the terms in $(\pi_2 - \pi_1) L_{s,t} - B^0_{s,t} \varepsilon_0$ have zero expectation (in other words the Lévy area of the Brownian motion has zero expectation). For a general driving path $Y_t \in \mathbb{R}^d$ of finite $2 + \varepsilon$-variation ($0 < \varepsilon < 1$) in the sense of [14] and [15], the 1-truncated scheme would not converge to the solution of the SDE. The use of the (at least) 2-truncated scheme would guarantee the convergence, but similar arguments as in Corollary 4.2 show that the scheme adjusted with the expected area, i.e. the ODE approach based on the vector field

$$
W_{s,t} := Y^0_{s,t} V_0 + \sum_{i=1}^{d} Y^{i}_{s,t} V_i + \frac{1}{2} \sum_{1 \leq i < j \leq d} \mathbb{E} \left[ Y^{(i,j)}_{s,t} - Y^{(j,i)}_{s,t} \right] [V_i, V_j]
$$

is sufficient for the convergence.
Remark 4.3. Note that the $\hat{W}^1$-based schemes explicitly or implicitly include the terms

$$\hat{X}_{t_i} B_{t_i,t_{i+1}}^{(j,j)} = \frac{1}{2} \hat{X}_{t_i} \left( B_{t_i,t_{i+1}}^j \right)^2$$

for $j = 1, \ldots, d$. On the global level, these terms form a non-vanishing contribution to the error, tending to the cross variation process i.e. the difference between the solution to SDE (3.1) and the solution to the SDE defined by the same coefficient functions but with Itô integrals.

5. Implementing high order schemes

The ODE approach based on the vector field $\hat{W}^1_{s,t} = \sum_{i=0}^d B^i_{s,t} V^i$ has global order 1 if $d = 1$ and 1/2 in general otherwise. To implement higher order schemes, the simulation of higher order Brownian iterated integrals is required. In this chapter, we introduce a method to overcome the difficulty of simulating higher order Brownian iterated integrals.

The key idea is the following. In general, the numerical schemes derived from $\pi^0_{s,t}$ make an error of certain order. We show that one can replace $\pi^0_{s,t}$ with another random Lie element which is in some sense close enough to $\pi^0_{s,t}$, and the ODE scheme derived from the replacement has the same global order as the original scheme. One particular way to find a good replacement to $\pi^0_{s,t}$, is to derive the log-signature of piece-wise linear approximations of the Brownian motion. This also can be interpreted as follows. Let us choose a fine scale together with the ODE approach based scheme corresponding to the scale and the vector field $\hat{W}^1_{s,t}$ where the exact solution to the local ODEs are considered. Globally, the scheme can be regarded as a non-autonomous ODE driven by piece-wise linear approximation of the Brownian paths on the chosen scale (equation (5.2)). Then we apply the ODE-based approach (just as it is derived in the earlier sections) to approximate the solution to this random non-autonomous ODE, i.e. we derive a high order approximation on a coarse scale of the low order approximation on the fine scale.

The benefit of this compound approximation is not straightforward, and at the end of the section, we give a brief analysis when it is worth implementing. Now, we continue with a rigorous derivation of the sketched method.

5.1. Piece-wise linear Brownian paths. We start our analysis with the tensor algebraic representation of linear paths.

Lemma 5.1. Let’s suppose that a path $\omega_t : [0, T] \to \mathbb{R} \oplus \mathbb{R}^d$ is linear, i.e.

$$\omega_t = (t, z_1 t, \ldots, z_d t)^T.$$  

Then the log-signature of $\omega$ is

$$\log(S_{0,T}(\omega)) = T \left( \varepsilon_0 + \sum_{i=1}^d z_i \varepsilon_i \right).$$

The multiplicative property of the signature (see Lemma 3.2) and Lemma 5.1 imply the following lemma.
Lemma 5.2. Given a piece-wise linear interpolation $B^P_t$ of a Brownian motion corresponding to the partition $P = (0, t_1, \ldots, t_{k-1}, T)$, i.e. the continuous concatenation of paths of the form
\[
\omega^i_0 = 0,
\omega^i_t = \left( t, B^1_{t_{i-1},t_i}, \ldots, B^d_{t_{i-1},t_i} \right)^T, \quad t \in (0, t_i - t_{i-1}]
\]
then
\[
(5.1) \quad L^P_s, T := \log(S^0_{s,T}(B^P)) = \log \left[ \exp \left( \sum_{j=0}^{d} B^j_{0,t_1} \varepsilon_j \right) \otimes \cdots \otimes \exp \left( \sum_{j=0}^{d} B^j_{t_{k-1},T} \varepsilon_j \right) \right]
\]
If the partition $P$ contains $k$ subintervals of equal length, we will use the notation $P_k$.

5.2. High order approximations based on piece-wise linear paths. In this section we prove, that if the numerical schemes are derived from $\pi_m L^P_{s,t}$ for a properly chosen $k$ instead of $\pi_m L^P_{s,t}$, the same global order of convergence is achieved.

Let’s denote the $i$th coordinate of $B^P_{s,t}$ by $B^i_{s,t}$. The following random ODE
\[
(5.2) \quad d \hat{\xi}_t = \sum_{i=1}^{d} V_i(\hat{\xi}_t) dB^i_{s,t}
\]
Since the equation (5.2) is a.s. driven by a path of finite length, we have a differentiation rule analogous to the Itô lemma
\[
(5.3) \quad df(\hat{\xi}_t) = \sum_{i=0}^{d} V_i f(\hat{\xi}_t) dB^i_{s,t}
\]
for any smooth function $f$.

We introduce the notation with analogy to section 2
\[
(5.4) \quad B^{J,P_{s,t}} := \int_{s<t_1<\cdots<t_k<t} \circ dB^{J_1,P_{s,t}} \circ \cdots \circ dB^{J_k,P_{s,t}}
\]
\[
(5.5) \quad B^{J,P_{s,t}}(Y) := \int_{s<t_1<\cdots<t_k<t} Y_{t_1} \circ dB^{J_1,P_{s,t}} \circ \cdots \circ dB^{J_k,P_{s,t}}
\]
for a continuous function $Y$. The fact that $B^{J,P_{s,t}}$ and $B^{J,P_{s,t}}(Y)$ are the lowest order ODE-approach based approximations of $B^J_{s,t}$ and $B^J_{s,t}(Y)$ respectively, is crucial for the proofs of the assertions in this section. A rigorous derivation of this fact can be found in [7].

Given the differentiation rule (5.3) one can derive the analogous version of the log-signature theorem.

Lemma 5.3. Let’s suppose that $t \in (0, 1)$, $\hat{\xi}_t$ is (a.s.) the unique solution for the random ODE (5.2) and $\hat{y}_s$ is the solution to the following random ODE
\[
(5.6) \quad \begin{cases}
\hat{y}_s &= \hat{W}(\hat{y}_s) ds \\
\hat{y}_0 &= \hat{\xi}_0
\end{cases}
\]
where \( \hat{W} = \Gamma \left[ \pi_m L_{0,t}^P \right] \). Then
\[
\hat{\xi}_t - \hat{y}_1 = \Gamma \left[ (\pi_{m+1} - \pi_m) L_{0,t}^P \right] (\hat{\xi}_0) + \mathcal{R}
\]
and there exist constants \( \mathcal{C}_1 \) and \( \mathcal{C}_2 \) not depending on \( t \) such that
\[
\mathbb{E} \left[ \| R \|_2 \right]^{1/2} < \mathcal{C}_1 t^{(m+2)/2} + \mathcal{C}_2 \left[ \frac{t}{k} \right]^\gamma,
\]
where \( \gamma = 1 \) if \( d = 1 \) and \( \gamma = 1/2 \) otherwise.

The proof of this lemma is very similar to the proof of the Log-signature theorem. The key difference lies in the derivation of the \( L^2 \) bound on the remainder term of the analogous stochastic-Taylor expansion of \( \hat{\xi} \). However, recalling that this remainder term is the lowest order ODE-approach based approximation of the remainder term of the stochastic-Taylor expansion of the exact \( \xi \), the existence of the constants \( \mathcal{C}_1 \) and \( \mathcal{C}_2 \) is implied.

**Remark 5.1.** The above lemma can be regarded as the probabilistic extension of [23]. From [23] one can derive an explicit formula for \( \Gamma \left[ \pi_m L_{0,t}^P \right] \).

By section 4.2, \( \hat{\xi}_t \) is also the extended ODE approach based \( m = 1 \)-truncated approximation of the solution for (3.1) on \([0, t]\) corresponding to the partition of \( k \) subintervals of equal length. Hence, we link Lemma 5.3 with Theorems 3.1 and 4.1 as follows.

**Theorem 5.1.** Let’s suppose that the conditions of Theorem 3.1 are satisfied and \( t \in (0, 1) \). Then there exists a positive integer \( k \) such that
\[
\xi_t - \hat{y}_1 = \Gamma \left[ (\pi_{m+1} - \pi_m) L_{0,t}^P \right] (\xi_0) + R
\]
where \( \hat{y}_1 \) is the solution to the random ODE defined in Lemma 5.3 and corresponding to the partition \( P_k \), and furthermore, there exists a constant \( C \) not depending on \( t \), such that
\[
\mathbb{E} \left[ \| R \|_2 \right]^{1/2} \leq C t^{(m+2)/2}.
\]

**Proof.** We split \( \xi_t - \hat{y}_1 \) into two terms as follows
\[
\xi_t - \hat{y}_1 = (\xi_t - \hat{\xi}_t) + (\hat{\xi}_t - \hat{y}_1)
\]
and regard the two terms separately.

Given that \( \hat{\xi}_t \) is an extended ODE approach based \( m = 1 \)-truncated approximation of \( \xi_t \) corresponding to the partition \( P_k \), by Example 4.2, there exists a constant \( C_1 \) such that
\[
\mathbb{E} \left[ \| \xi_t - \hat{\xi}_t \|_2 \right]^{1/2} \leq C_1 \left[ \frac{t}{k} \right]^\gamma
\]
where \( \gamma = 1 \) if \( d = 1 \) and \( \gamma = 1/2 \) otherwise.

Lemma 5.3, equation (5.8) imply the following representation
\[
\xi_t - \hat{y}_1 = \xi_t - \hat{\xi}_t + \Gamma \left[ (\pi_{m+1} - \pi_m) L_{0,t}^P \right] (\xi_0) + \mathcal{R}
\]
where \( \mathcal{R} \) is defined in Lemma 5.3.
Define
\[ R := \xi_t - \hat{\xi}_t + \left[ (\pi_{m+1} - \pi_m)L_{0,t}^{P_k} \right] (\hat{\xi}_0) - \Gamma \left[ (\pi_{m+1} - \pi_m) L_{0,t} \right](\xi_0) + \mathcal{R} \]

Let's rewrite \( \Gamma \left[ (\pi_{m+1} - \pi_m)L_{0,t}^{P_k} \right](\hat{\xi}_0) \) as follows

\[ \Gamma \left[ (\pi_{m+1} - \pi_m)L_{0,t}^{P_k} \right](\hat{\xi}_0) = \sum_{J \in A} \|J\| = m+1 \left( \sum_{J_1, \ldots, J_l \in J, i \leq |J|} a_{J_1, \ldots, J_l} B_{0,t}^{J_1, P_k} \cdots B_{0,t}^{J_l, P_k} \right) \Gamma(\varepsilon_J)(\hat{\xi}_0) \tag{5.10} \]

with deterministic real coefficients \( a_{J_1, \ldots, J_l} \) depending only on the multi-indices \( J_1, \ldots, J_l \).

Since \( B_{0,t}^{J, P_k} \) is the lowest order ODE approach based approximation of \( B_{0,t}^J \) for each multi-index \( J \), therefore there exists a constant \( C \) not depending on \( t \) or \( k \) such that

\[ \mathbb{E} \|R\|^2 \leq \mathbb{E} \|\mathcal{R}\|^2 + C \left[ \frac{t}{\kappa} \right]^\gamma \tag{5.11} \]

The proof is complete by choosing a high enough \( k \) satisfying

\[ (C + \mathcal{C}_2) \left[ \frac{t}{\kappa} \right]^\gamma \leq \overline{C}_1 t^{(m+2)/2} \tag{5.12} \]

where \( \overline{C}_1 \) and \( \mathcal{C}_2 \) are the constants introduced in Lemma 5.3.

Theorem 5.1 shows, that for high enough \( k \), \( \hat{y}_1 \) and \( y_1 \) (corresponding to the \( m \)-truncated log signature) are numerically equivalent (local) approximations of \( \xi_t \). However, the left-hand side of (5.12) should not be much smaller than the right-hand side, because for high values of \( k \) the computation of \( \Gamma \left[ \pi_m L_{0,t}^{P_k} \right] \) is less tractable, i.e.

\[ k \approx \left[ \frac{C}{C_1 t^{(m+2)/2 - \gamma}} \right]^{-1/\gamma} \tag{5.13} \]

is the recommended order of magnitude for \( k \).

5.3. Practical considerations. The implementation of the high order schemes based on piece-wise linear approximation of the Brownian paths on a fine sub-scale requires the repeated computation of \( \Gamma \left[ \pi_m L_{0,t}^{P_k} \right] \), where \( L_{0,t}^{P_k} \) is given by (5.1). To work out the referred formula, one could use the Campbell-Baker-Hausdorff formula (ref.: [23]) at each discretization step, but this would not be efficient.

Furthermore, the principle given by (5.12) and (5.13) implies that shorter time intervals require higher values of \( k \) to ensure the high order convergence of our scheme. However, for higher values of \( k \) the computation of (5.1) increases rapidly and one would lose the linear growth of computational expense.

Truncating (5.1), we get the following assertion.
Lemma 5.4. The log-signature (5.1) written in a Lie basis \( \{ \ell_i, i \in \mathbb{N} \} \) is of the form

\[
\pi_m L_{0,t}^\mathcal{P} = \sum_{h, \pi_m \ell_h = \ell_h} p_h \ell_h
\]

where each \( p_h \) is a polynomial of the variables \( \{ B_{j,t_i}^i, j = 0, \ldots, d, i = 1, \ldots, k \} \).

The reader is referred to the monograph [21] by Reutenauer for details on Lie bases.

Example 5.1. Let’s assume that a piece-wise linear path is a concatenation of a number of \( k \) pieces with log-signature of the form:

\[
\sum_{r=0}^d b_{r,i} \varepsilon_r, \ i = 1, \ldots, k.
\]

Then the log-signature of the concatenated path is

\[
\left( \sum_{i=1}^k \sum_{r=0}^d b_{r,i} \right) \varepsilon_r + \left( \frac{1}{2} \sum_{1 \leq i < j \leq k} \sum_{r,s=1}^d b_{r,i} b_{s,j} \right) [\varepsilon_r, \varepsilon_s] + \ldots
\]

The terms in the parentheses are polynomials of the coefficients \( b_{r,i}, i = 1, \ldots, k, \ r = 0, \ldots, d. \)

Note that the structure of (5.14) does not change if \( k \) is increased but the polynomials \( p_i \) have more variables. Hence to preserve efficiency, one can pre-compute either this formula or the polynomial coefficients therein for a number of independent Brownian paths. The pre-computation might take some time, however it only has to be done once. After all, the reuse of the pre-computed values results in a fast and high order numerical algorithm.

Since the formula (5.14) does not depend on the SDE but on the dimension and the choice of \( k \), one could create a universal database of pre-computed coefficients usable for many SDEs.

We draw the reader’s attention to the fact that software packages are available for the algebraic computations sketched above, in particular: the \texttt{libalgebra} library created by the \textit{Computational Rough Paths} project in C++. See the project’s website \url{http://coropa.sourceforge.net/} for further details.

5.4. A note on the computational expense. The implementation of the high order scheme presented in this section is after all a high order approximation of a lowest order scheme corresponding to a discretization on a finer sub-scale. One might ask if one can ever benefit from working out the high order scheme instead of just solving the lowest order scheme on the fine scale. Let’s do a brief cost-analysis to answer the question.

Let’s assume that the computational cost of evaluating one step corresponding to the lowest order scheme is \( E_1 \) and the cost of evaluating one step of an \( m \)-truncated ODE based scheme is \( E_2 \). In general \( E_1 \) is much smaller than \( E_2 \). If we apply the lowest order scheme on an interval with a \( k \)-substep fine scale, then the cost is \( kE_1 \). As long as \( kE_1 \) is smaller than \( E_2 \), the lowest order scheme is recommended. However, if one needs a more accurate approximation and wants
to choose a finer scale, then $k$ will increase according to (5.13) and eventually the implementation of the higher order scheme will become more efficient than the corresponding low order scheme.

6. Numerical examples

In this section, some numerical examples for the ODE approach are presented. We demonstrate the efficiency of different ODE solvers and the high order methods. The primary aim is to estimate and compare the order of strong convergence, however we also present some weak approximation results.

6.1. Estimating the order of convergence. For the estimation of the convergence order, we apply some ideas from [22]. Let’s review some basic results.

Definition 6.1. Let’s regard a discretization method generating $\hat{X}_T$ as an approximation of $\xi_T$. If the discretization scheme is based on partitioning $[0, T]$ into subintervals each of length $\Delta t$, then this particular case, the resulting approximation is denoted by $\hat{X}_T(\Delta t)$.

Lemma 6.1. Let’s suppose that a discretization method $\hat{X}_T$ approximates $\xi_T$, i.e.

$$E \left[ \| \xi_T - \hat{X}_T(\Delta t) \|^2 \right]^{1/2} \to 0 \quad \text{as } \Delta t \to 0$$

Furthermore, let’s suppose the existence of constants $C_2$ and $\gamma$ not depending on $\Delta t$ such that for $0 < \Delta t < 1$

$$E \left[ \| \hat{X}_T(\Delta t/2) - \hat{X}_T(\Delta t) \|^2 \right]^{1/2} \leq C_2 \Delta t^\gamma$$

is satisfied. Then for $0 < \Delta t < 1$

$$E \left[ \| \xi_T - \hat{X}_T(\Delta t) \|^2 \right]^{1/2} \leq C_2 \Delta t^\gamma \frac{1}{1 - (1/2)^\gamma}$$

The proof of the lemma is based on the triangle inequality.

We apply Lemma 6.1 to estimate the order of convergence $\gamma$ and an upper bound on the strong approximation error as follows. Estimating the expected value on the left hand side of (6.1) for a sequence of step lengths $T/2^k$ for $k = l, l + 1, \ldots$ where $T/2^l < 1$ and fitting a log-regression, one can estimate $C_2$ and $\gamma$. If $C_2$ is estimated, then by (6.2) we get an estimation for an upper bound of the error at step size $\Delta t$. When $\gamma$ is known, we only estimate $C_2$.

As it was demonstrated in [22], one can observe the expected value on the left hand side of (6.1) very accurately, if the schemes corresponding to the different step sizes $\Delta t$ and $\Delta t/2$ are run on the same Brownian paths. In that case, the random variables $\hat{X}_T(\Delta t)$ and $\hat{X}_T(\Delta t/2)$ are correlated and a simple Monte Carlo method results in a low variance and unbiased estimation of the expected value of their distance. Running schemes corresponding to different step lengths on the same paths can be achieved by using Lévy’s construction of the Brownian paths (see e.g. [5] for details).

The idea of running schemes on the same Brownian paths is also useful when comparing two different discretization methods generating $\hat{X}_T$ and $\overline{X}_T$ respectively.
as approximations of \( \xi_T \). The high correlation between \( \hat{X}_T \) and \( X_T \) results in a low variance for the Monte Carlo estimation of

\[
E \left[ \| X_T - \hat{X}_T \|^2 \right]^{1/2}
\]

This comparison can be further applied as the triangle inequality implies.

**Lemma 6.2.** Let two discretization methods producing \( \hat{X}_T \) and \( X_T \) respectively as approximations of \( \xi_T \) be given. Let’s suppose that there exists a constant \( C \) such that

\[
E \left[ \| \xi_T - X_T \|^2 \right]^{1/2} < C
\]

then

\[
E \left[ \| \xi_T - \hat{X}_T \|^2 \right]^{1/2} \geq E \left[ \| \hat{X}_T - X_T \|^2 \right]^{1/2} - C
\]

Lemma 6.2 implies, that if one can estimate a small enough upper bound \( C \) on the global error of \( X_T \), then (6.4) gives an accurate estimation for the lower bound on the error of \( \hat{X}_T \).

In the following sections, for different SDEs and discretization methods we ran simple Monte Carlo simulation based estimations of the following quantities

(i) the estimated upper bound of the global error, i.e. (6.2)
(ii) the \( L^2 \) distance of different schemes, i.e. (6.3)
(iii) in some cases the estimated lower bound on the global error, i.e. (6.4)
(iv) and \( E [\hat{X}_T(\Delta T)] \)

for different step sizes. Each method was run on \( 10^6 \) paths.

Since each Monte Carlo simulation is based on sampling, the resulted estimation is a random variable with positive variance. Using this variance, when estimating (iv), we fit a 99\% confidence interval centred at the resulting realization of the random variable. The length of these confidence intervals is proportional to the square root of the number of runs. In case of (i), (ii) and (iii), the calculated confidence intervals were very small.

### 6.2. First order approximation of the CIR process.

Firstly, we regarded simple SDEs, driven by one dimensional Brownian motion and tested the first order ODE approach with different ODE solvers as well as comparing it to the Euler-Maruyama scheme. The test results, presented in Figures 2 and 3, are the weak and strong approximation results respectively for the CIR process, i.e. in the Itô form

\[
dr_t = a(b - r)dt + \sigma \sqrt{r}dB
\]

where \( a, b \) and \( \sigma \) are positive constants satisfying \( ab/2 > \sigma^2 \), which ensures that \( r_t \) is a.s. positive.

We implemented the ODE approach with three different ODE solvers, namely the predictor-corrector, Runge-Kutta order 4 and splitting. By splitting we mean the ODE solver in which \( W \) is written as \( W := W_1 + W_2 \) where

\[
W_1(x) := B^0 [a(b - x) - \sigma^2/4]
\]

\[
W_2(x) := B^1 \sigma \sqrt{x}
\]
and at each step (starting at $\hat{X}_{t_i}$), three ODEs are solved as follows

$$x_1 := \exp\left[\frac{1}{2}W_1\right](\hat{X}_{t_i})$$
$$x_2 := \exp\left[W_2\right](x_1)$$
$$\hat{X}_{t_{i+1}} := \exp\left[\frac{1}{2}W_1\right](x_2)$$

In case of the CIR SDE, the exact solution for the ODEs appearing in the splitting is known, and the computation is very fast. The splitting method is recommended in for example [20] and some nice weak approximation properties of it are presented as well.

The first order ODE approach implemented with the predictor-corrector ODE solver is referred to as the Heun scheme.

**Remark 6.1.** Note, that despite the guaranteed positivity of $r_t$, all of the tested schemes except the Splitting version can result in negative interest rates. In case of the ODE approach based method, the exact solution to the derived random ODEs preserves the positivity, but their numerical approximations might not. One possible way to overcome this difficulty, is to adaptively reduce the ODE numerical solver’s step size when the solver results in negative interest rates. In case of the splitting, the chosen combination of parameters guarantees positive solutions and no extra care is required. The Euler-Maruyama method is fixed by taking a positive value of the resulting interest rate at each step. We refer the reader to [13] for a review of simulation schemes approximating the CIR process.

Figure 2. Weak approximation results (CIR)

In Figure 2, the weak approximation results, i.e. the confidence intervals corresponding to Monte Carlo estimations for $\mathbb{E}[\hat{X}_T]$, are presented for different schemes implemented with different numbers of steps. The horizontal dashed line is the exact value of $\mathbb{E}[\xi_T]$. Given the number of runs ($10^6$) the variance of the evaluated estimation was relatively large, and at the weak level one cannot make a difference between the Runge Kutta 4 and the splitting versions. The weak error of the Euler method at 64 steps seems smaller than the weak error of the Runge Kutta 4,
but the calculated values are samples from random variables with positive variance. Note that the Euler-Maruyama method results in a first order weak approximation, i.e. at the weak level there is no difference in the convergence order of the tested methods.

<table>
<thead>
<tr>
<th>Number of steps</th>
<th>Estimated error</th>
<th>Euler-Maruyama</th>
<th>Euler-M. lower bound</th>
<th>Heun</th>
<th>Runge Kutta 4</th>
<th>Splitting</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$10^{-6}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>$10^{-5}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>$10^{-4}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>$10^{-3}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>$10^{-2}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>32</td>
<td>$10^{-1}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Figure 3. Estimated error, comparison of solvers (CIR)**

In Figure 3, the accuracy of the path-wise approximations is presented. The graph on the left-hand side presents the estimated upper bound on the error based on (6.2). In case of the Euler method, we could apply (6.4) and a lower bound on the error is estimated. This lower bound demonstrates, that despite the nice weak approximation properties, the Euler-Maruyama scheme is less accurate in the path-wise approximations when compared to the other schemes. The estimated order of convergence (i.e. the slope of the curves) are as expected. In case of the Euler-Maruyama scheme it’s close to $1/2$ and in case of the ODE approach based methods it is close to $1$.

The graph on the right-hand side of Figure 3 compares each method with the Runge Kutta 4 version, estimating the $L^2$-distance based on (6.3). Note that the $L^2$ distance of the splitting version from the Runge Kutta 4 based method has an order of magnitude $10^{-6} - 10^{-7}$, whereas the estimated global error of both methods has order $10^{-5}$ and so numerically the two versions are equivalent. However the splitting is a bit faster than the Heun scheme and more than twice as faster than Runge Kutta 4, so in this particular case the Splitting version is recommended.

Each ODE solver tested here can be easily extended to handle the term $\int_0^t r_u \, du$. The simulation of this term is required when pricing or hedging bonds and derivatives of bonds. In case of the Splitting, the exact solution to the new ODEs appearing in this extension are known.

**6.3. Second order approximation implemented.** In this section we present some numerical results of some tests run with the second order ODE approach
based method. The SDE chosen here is less meaningful in finance. However due to its’ nice properties it is proved to be a useful test case.

The SDE is as follows

\[
\begin{align*}
\frac{dx_1^1}{t} &= \sin(x_1^1) \circ dB^1 \\
\frac{dx_1^2}{t} &= \sin(x_1^1) dt
\end{align*}
\]

The left-hand side of Figure 4 presents the weak approximation results in the form of confidence intervals. According to the rule (5.12), for the 1-step second order version we used a linear interpolation of the Brownian motion on a two step sub-scale, i.e. \( k = 2 \). In case of the 2-step version \( k = 4 \) was chosen whereas the 4-step version was run with \( k = 8 \).

![Figure 4. Implementing high order approximations](image)

The estimated upper bound on the strong global error is presented on the right-hand side of Figure 4. The second order method, implemented as described in section 5 has an estimated order close to 2.

We analysed the computational expense for this particular SDE as described in section 5.4. In this case \( E_2/E_1 \approx 3/2 \), so the second order scheme is computationally more efficient than the lowest order at every step. For higher order schemes and in higher dimensions, the lower step high order versions might be relatively less efficient.

7. Concluding remarks

The approach presented by this paper turned out to be very useful in the case of strong approximations as well as in the case of weak approximations. The ideas presented here can be applied to construct weak approximations under reasonably general conditions (see [10], [11] and [12]). The concept of Cubature on Wiener space ([16]) led to the construction of finitely supported measures on the truncated Lie algebra converging to the exact distribution of the Brownian log-signature and
resulting in high order weak approximations. The application of these ideas in computational finance is described in [7].

The theorems of this paper ensuring the global accuracy are based on global and in some cases rather strong conditions. The focus of a further work could be to explore how these conditions could be weakened and how one could replace the global conditions with local ones.

References


Mathematical Institute, Oxford University, 24-29 St Giles’, Oxford, OX1 3LB, UK
E-mail address: gyurko@maths.ox.ac.uk

Mathematical Institute, Oxford University, 24-29 St Giles’, Oxford, OX1 3LB, UK
E-mail address: tlyons@maths.ox.ac.uk
Hedge Funds

Luis A. Seco, Fangyuan Chen

Abstract. This paper describes the investment universe of hedge funds from a perspective which is at the same time mathematical in nature and practical in its objectives. It addresses the investment opportunities that hedge funds pursue, the mathematical techniques to understand their performance, the portfolio construction tools needed to construct investments, and risk methodologies useful in assessing performance.

1. What is a Hedge Fund?

The best way to understand it is to actually construct one. Let’s do it.

1.1. Example: the snow swap. Canadian winters are extreme: cold and snow are a fact of everyday life. Canada spends over $1Bn every year removing snow. As one example, consider the city of Montreal. The city spends over $50M every year removing snow, about 3% of its total budget. It does that through a fixed-price contract agreement with a third party, and in the contract, the “snow season” is specified during which the third party is responsible for removing snow in the city. In the case of Montreal, this period starts on November 15 and ends in April 15. During this defined period, the city’s exposure to snow removal costs are to a large degree predictable (see [FLM] for a detailed account.) However, snow precipitation outside of this period can become very costly: it is outside of the contractual arrangement, and the city may incur expenses which may, on a relative basis, exceed the ones during the snow season.

We say cities, such as Montreal, are exposed to snow financial risk. However, snow financial risk also affects other corporations, such as ski resorts. For them, the snow financial risk is opposite: low precipitation during the late part of the fall or early spring will yield operational losses compared to years when snow fall is ample early in the fall or late into the spring.

Sometime ago, a proposal was launched to partially mitigate this: a snow swap. In this, a city will pay a premium to a dealer when snow is scarce outside the snow season, and receive a premium if snow appears. Similarly, a ski resort will receive payments if snow is scarce and will pay if snow is plentiful. The dealer acts as
intermediary in the swap, and collects a commission for its services. The dealer has no risk exposure to snow precipitation because it is exchanging offsetting payments between the two parties.

The snow swap did not succeed, however, because there was no agreement between the city and the ski resort as to where the measurements for snow precipitation should be taken. Specifically, ski resorts are mostly located in suburban areas while majority of the city is located in urban areas. Therefore, ski resorts would like to measure the snow precipitation at suburban areas, while the city would like it be at urban areas. The snow amount in suburban and urban areas do not always agree, giving rise to the geographical spread of snow precipitation. The snow financial risk seemed to be solved by the snow swap, but the geographical spread risk could not be absorbed by anyone.

Let us consider the hypothetical following proposition: a group of investors (a fund) gets together, puts up some money upfront as collateral, and decides to take the geographical spread risk. It will pay the city in the case of out-of-season snow falls in the city, and will pay the ski resort in case of no out-of-season snow falls at the resort (See Table 1). By contrast, it will receive payments from both if the opposite occurs. With an nominal payment of $1M, and an nominal fee of 10% ($100,000), cash flows for the fund is shown in Table 2. The difference with the previous, unsuccessful snow swap is that in this case, both the city as well as the ski resort get to measure the snow precipitation at the place of their choice, with the fund taking the geographical risk.

<table>
<thead>
<tr>
<th>Payments</th>
<th>Snow</th>
<th>NoSnow</th>
</tr>
</thead>
<tbody>
<tr>
<td>City</td>
<td>$ - 1M</td>
<td>$1M</td>
</tr>
<tr>
<td>Ski resort</td>
<td>$1M</td>
<td>$ - 1M</td>
</tr>
<tr>
<td>Fee</td>
<td>$100,000</td>
<td>$100,000</td>
</tr>
</tbody>
</table>

To move ahead with our example, let us assume the snow events in both places are correlated at 50%, and the fund will charge the same $100,000 fee from both counter-parties for its risk: this means that the cash-flows for the fund will be:

<table>
<thead>
<tr>
<th>Event</th>
<th>CashFlow</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Offset payments</td>
<td>$200,000</td>
<td>75%</td>
</tr>
<tr>
<td>Pays both</td>
<td>$-1,800,000</td>
<td>12.5%</td>
</tr>
<tr>
<td>Receives from both</td>
<td>$2,200,000</td>
<td>12.5%</td>
</tr>
</tbody>
</table>

1.1.1. *Earning perspective of the fund.* To get an idea of the quality of these funds, we look at two basic properties: expected return and standard deviation. Suppose the investors need to put down the maximum amount that the fund needs to pay out, $2M, in the case of snow falls in the city, but not in the ski resort. Then
the investment required to set up this swap is $2M, and expected return based on Table 2 is $200,000. That is a return of 10%, comparable to an investment in the stock market. The standard deviation, however, is at 50%, which is more or less comparable to a game of poker. From an investment viewpoint, this is not a very good proposition, as the risk is too high for the expected return. Things become more interesting if the fund decided to do similar swaps in other cities. If 100 independent swaps are considered, for a total of $200M invested, the expected return continues to be 10%, but the standard deviation, as a measure of risk, now drops to 5%, because of the diversification. As an investment, this is now better than investing in the stock market and the fund becomes attractive. We have the basic concept to start a hedge fund now.

1.1.2. Financing the investment. In our snow fund, we raised $200M to post as collateral for 100 different swap agreements. This was to give rise to an expected return of 10% ($20M) for the period (6 months), with a standard deviation of 5%. Note that in calculating our cash-flows, we have neglected the fact that the collateral ($200M) was not to be used except as a guarantee to both counter-parties. With $200M, our fund would be able to honor its payment obligations even when all deals may turn against the fund. In other words, the collateral is there just to enable the fund to have the right credit rating for the deal. The fund would obtain a rating of AAA, the best possible. However, there is no reason to hold the $200M in cash, one could easily invest them in T-Bills (short term interest notes issued by the government of the United States), and hence earn LIBOR, the on-going risk-free interest rate. In this way, our return will be LIBOR+10%, with a standard deviation almost unchanged.

1.1.3. Leverage. As mentioned above, we need to raise $200M to start the 100 snow swap agreements. One direct way is raising the whole amount from investors. Alternatively, we can borrow part of the amount, which leads to what we call leverage. Leverage affects the returns to fund investors, who are effectively shareholders of our hedge fund. We illustrate this point in more detail.

Assume now we can borrow part of the investment needed at a fixed interest rate of 4%. In Table 3, we show the calculations related to different levels of leverage.

<table>
<thead>
<tr>
<th>Leverage</th>
<th>0%</th>
<th>20%</th>
<th>50%</th>
<th>80%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Investment</td>
<td>$200M</td>
<td>$200M</td>
<td>$200M</td>
<td>$200M</td>
</tr>
<tr>
<td>Borrowed Amount</td>
<td>0</td>
<td>$40M</td>
<td>$100M</td>
<td>$160M</td>
</tr>
<tr>
<td>Equity</td>
<td>$200M</td>
<td>$160M</td>
<td>$100M</td>
<td>$40M</td>
</tr>
<tr>
<td>Interest rate</td>
<td>4%</td>
<td>4%</td>
<td>4%</td>
<td>4%</td>
</tr>
<tr>
<td>Expected Cash Flow</td>
<td>$2M</td>
<td>$2M</td>
<td>$2M</td>
<td>$2M</td>
</tr>
<tr>
<td>Interest Payment</td>
<td>0</td>
<td>$1.6M</td>
<td>$4M</td>
<td>$6.4M</td>
</tr>
<tr>
<td>Expected return of equity</td>
<td>5%</td>
<td>11.5%</td>
<td>16%</td>
<td>34%</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>5%</td>
<td>6.18%</td>
<td>10%</td>
<td>25%</td>
</tr>
</tbody>
</table>
We show the leverage-return and leverage-standard deviation relationship in Figure 1 and Figure 2. From both the table and the figures, increasing leverage increases the expected return of equity investment, but coming with it equally higher risk associated, represented by standard deviation of returns. This is the well-known feature of leverage usage in investments.

**Figure 1.** Expected return on equity with different leverage levels

**Figure 2.** Standard deviation of return with different leverage levels
1.1.4. Discussion of Credit Risk. In our discussion above, we have assumed that we need to finance $200M to serve as collateral, which is enough to pay for all obligations even if the fund needs to pay both sides in all 100 snow swap contracts. This would give us a Aaa credit rating (as in Moody’s credit rating system).

However, in [Figure refrating], we see that the investors own the hedge fund, which has a Aaa credit rating. However, the cities and ski resorts are mostly likely to have lower credit ratings. For example, city of Montreal had a A1 credit rating and was upgraded to Aa2 in 2008. In such a case, the fund faces credit risk. More specifically, if there is little snow in the city and plenty of snow in the ski resort, the fund is supposed to receive payments from both sides. However, one or both counter-parties might default on their obligations.

![Figure 3. Illustration of Credit Risk](image)

From [Figure 3], we can see that it is not necessary for our fund to keep the best credit rating, since the counter-parties have lower ratings. The fund can put down less than $200M as collateral. For example, we can raise only $120M and have a credit rating of Aa2, the same as the city. This reduction in financing needed compensates for the fact that the fund will be bearing credit risk from possible defaults of counter-parties.

1.2. Investment Structure of Hedge Funds. The previous investment idea makes sense. Imagine we have already come to agreements with our counterparties (the cities and sky resorts) and are ready to create our portfolio of snow swaps. If we have the necessary funding, all $200M, we can simply sign the agreements and wait for our investment to succeed. Most of the time, however, we don’t have the funds, just the idea, and therefore we need to finance it. There will be two things that we will have to do.

1. We need to first raise the money. It turns out that most jurisdictions limit our ability to raise money. We may need to be registered as an investment advisor, we may have to register the investment terms with a government regulator, and we may have limitations as to which type of institutions or individuals we can approach, etc. In legal parlance, we need to create a fund. A fund is the investment vehicle that allows us to raise money from third parties to invest in a common portfolio, and the fund will be owned by the investors, not us.

If our fund were dealing with investment in stocks or bonds, we could do all that, with enough money and legal help. If we are planning to
invest in exotic financial instruments such as snow swaps, we are unable to do so. Many jurisdictions will not allow us to do that. We may have to constitute our fund in places such as the Cayman Islands or Bermuda. They are jurisdictions with a sound legal system but with less regulatory hurdles than the US, Canada or Europe.

(2) The fund will collect the investor’s money, and become the legal counter-party to the snow swaps. There is another thing we should do: managing the fund. Without a management structure, the fund will be like a ship without a captain. The fund will have money but won’t have the ability to do anything with it. We need to create the management company of the fund and implement our investment ideas. The fund manager gets paid for managing and improving the performance of the fund, and the manager’s salary and bonus are often closely tied to the performance.

Understanding the separation of the fund and its management is very important. The term hedge fund is often times used to refer to the fund or the management company, without realizing that they are two completely different entities.

Investors invest in our fund, and we create a legal structure for the fund, consisting of three parts, as shown in [Figure 4].

**Figure 4. Hedge Fund Investment Structure**

- Management
  This is mentioned above as one of the two key procedures to get the fund started.
- Prime broker
  The hedge fund needs a prime broker to executive trades for us, such as long/short stocks, purchase bonds, and investing in other financial instruments.
- Administration
  The separation of these three roles can be important in ensuring performance of the fund and preventing fraud. If the management firm acts as both the fund manager and the prime broker, it can report to the investors profitable trades that
HEDGE FUNDS

are not actually executed, and therefore very attractive returns that are not real. Separation of duties imposes some supervision and is a critical part of internal control in preventing fraud.

In addition to these three components, hedge funds also hire auditors to audit its accounting practices. The work of a hedge fund auditor includes verifying valuation methodologies stated in the Offering Memorandum, review financial statements prepared by management, and check transaction records as well as profit/loss statements.

The largest fraud committed by a single person, the Madoff Scandal, is the best example to show the importance of such separation. The Bernard L. Madoff Investment Securities LLC documented fake trading activities for his investment advisory firm through its own broker-dealer. The investigators found later that there is no evidence or record showing these trades were executed. The company also had its own back-room accounting firm that claimed to have audited all the performance reports. The accounting firm in fact has only three employees and is impossible to handle the audit work of a multi-billion financial business. The three roles including management, prime broker, and audit were all within the Madoff Investment Security LLC and managed by Madoff himself, without any third-party supervision.

1.3. Features of hedge funds. Our snow fund example embodies some of the main features of what constitutes a hedge fund:

(1) A hedge fund is a private investment partnership; the hedge fund manager sets up a fund -the hedge fund- as a legal entity that allows the manager to attract money from investors and invest it according to certain rules.

(2) The investments that the fund can do are not regulated by government agencies; instead, they are regulated by a private document, the Offering Memorandum, which the manager creates as part of the legal structure of the fund, and the investors accept as part of their investment. This is a key document, as it sets the boundaries of what the manager can do with the investor’s money. For example, in our snow fund, it may allow the manager to enter into snow swaps with cities and ski resorts, but it may not allow the manager to invest in the stock market. However, it may allow the manager to invest in bonds, since the manager may prefer to obtain interest from the fund’s assets while they are used as collateral for the swaps. This lack of regulation must not be confused with the fund’s need to comply with investment law of each jurisdiction it participates, which of course must always be upheld by the manager; the hedge fund manager bears the full liability of compliance with securities and other laws.

(3) A hedge fund seeks return niches by taking risks, which they may hedge or diversify away (or not). Hedge funds are often set up to exploit what is perceived as a market inefficiency by the manager: a possibility to make money in excess of the interest rate with low risk. This inefficiency can arise from the profitability of snow swaps, from the inability of the market to price certain bonds correctly, from opportunities created by companies on the brink of bankruptcy, or companies with potential for mergers or acquisitions, etc. A hedge fund is usually focused on one type of inefficiency, which defines their trading style and may be reflected on
their Offering Memorandum. Some larger hedge funds may include a variety of styles into their trading portfolio.

(4) They seek returns independent of market movements. For example, the performance of our snow fund is likely to be independent of whether stock markets go up or down, and although they may have small correlation with bond yields—perhaps due to the fact that the collateral is invested in bonds—the main driver for their return—the profitability of snow swaps—has no relationship to interest rates. Some of the first hedge funds were created by traditional stock pickers in the 1950's who eliminated their exposure to the broad stock market by adding a short stock index position\footnote{A short stock position refers to an investment in a stock which consists of a negative amount of shares; the short investor therefore profits the price of the investment falls and loses when the price of the portfolio drops} to their portfolio; in this way, the return is produced if the stocks selected outperform the market, which can happen even when the market is down.

(5) Hedge funds Net Asset Value (NAV) is reported monthly. This is a key number, which reflects the value the investors own. As the investors incur into gains, this number increases, and it decreases when investor’s lose money. However, this number can also increase as more investors join the fund, and decrease as they redeem their investments. This number is calculated by adding all the funds assets and removing the value of the liabilities, which include the fees payable to the fund manager and other costs. In evaluating the fund assets, the snow swap, we need a method or a model to value the snow swaps we have constructed. There is no ready valuation formulas, and the cash flow for the swap is not deterministic. The cash flows can not be foreseen accurately because our counter-parties are default-prone and not risk-free. When some counter-parties become more likely to default, the value of our swap would decrease. These factors all contribute to difficulty of reporting NAV monthly.

(6) The fund management company works for the fund, and therefore the fund pays the management company. The fees vary widely from one fund to another, and are stipulated in the Offering Memorandum, but they have two components: one is a flat management fee, which may vary from 1% to 2% of the NAV, and is paid whether the fund makes money or not. The second one is a performance fee, which often ranges from 10% to 20% of the net gains of the fund. Normally, the management fee is paid monthly or quarterly, while the performance fee is paid annually.

1.4. Fund Structure. Investors participate in the fund through the purchase of fund shares, which are analog to the shares issued by companies. If we denote the fund’s Net Asset Value (NAV) by \( N \) and if the fund has \( n \) shares outstanding, the price per share \( S \) is equal to

\[
S = \frac{N}{n}
\]

This number will increase as investors gain, and decrease as they lose. Note that this definition is a relative one, that is the assets can grow because of new investors join the fund, or decrease because existing investors redeem out of the fund, while the price per share remain unchanged. In this regard, \( S \) changes only as the underlying asset value changes relative to investor participation. From an accounting...
viewpoint, the share price and number of outstanding shares contains most of the required information of the fund.

As we saw earlier, this number is reported monthly. Hence, it gives rise to a time series of monthly observations, which we will denote by $S_1, S_2, \ldots, S_i$, where the subindex $i$ refers to a month.

1.5. Types of Hedge Fund Investments. In the early times of hedge funds, investors used to invest in those funds directly. Nowadays, investors are often times invested in hedge funds, not directly, but through a variety of investment products, which we now briefly review.

1.5.1. Fund of funds. A management company, independent of the management company of hedge funds, but with very similar characteristics, creates a fund whose assets are invested in a portfolio of hedge funds. They are second-level hedge fund investments, where the investor acquired shares in a single fund, whose assets are then used to invest in a series of hedge funds.

The advantages of such investments are:

- Diversification
- Outsourcing to the fund-of-funds management company of proper due investment diligence tasks, analytics and other processes related to hedge fund investments
- Access to hedge funds which may be closed to other investors, or specially favorable investment terms with the underlying hedge funds

Among the disadvantages, we have

- Higher fees: the investor will have to pay the management and performance fees of the fund-of-funds, which are over and above the ones for the underlying hedge funds
- Increased default risk: if an investor invests in a portfolio of hedge funds, and one of them bankrupts, the losses are proportional to the participation in the defaulted fund. But if the fund-of-funds defaults (for example, due to fraud at the fund-of-funds level) then all the assets are at risk.

1.5.2. Hedge Fund Indices. Hedge fund indices are composite portfolios designed to reflect hedge fund industry performance. The indices are constructed to be benchmarks that are representative of the hedge fund industry performance. Methodologies are used to select constituent hedge funds and produced weighted returns. For example, the HFRX Indices methodology combines both quantitative and qualitative processes. Statistical analysis is used to replicate statistical properties of each hedge fund, and this include correlation analysis, optimization, and Monte Carlo simulations. These methods are used to generate the highest statistical likelihood of producing a return series that is most representative of the composite hedge fund returns. Qualitative criteria such as transparency of the fund and manager’s due diligence are also considered before a fund can be included.

Hedge fund indices are often classified by the strategy employed or regional focus. Examples of indices by strategy are HFRX Equity Hedge Index, HFRX Equity Neutral Index, HFRX Convertible Arbitrage Index, HFRX Distressed Securities Index, etc. Indices by regional focus include HFRX Asia Composite Hedge Fund Index, HFRX Total Emerging Market Index, and HFRX North America Index, etc.
Hedge Fund Indices, in contrast with stock or bond indices, can be investable, or non-investable. By including hedge funds that are open to investors, the hedge fund index can be turned into an investable product, and they often are. Examples of such indices are the HFRX indices, managed by HFR (Hedge Fund Research Inc.), a Chicago-based management firm. However, hedge fund indices may contain funds which are not open to new investors, in which case the indices are not investable and cannot even be replicated. Example of such indices are the HFRI indices, produced also by HFR, and older than their siblings HFRX indices. For an index to be investable, proper legal structures must be set up, such as a fund structure. This is another reason why many firms who publish hedge fund indices, whether closed or open to new investors, produce exclusively non-investible indices. An example of these would be the indices provided by Barclays.

1.5.3. Structured Products. Structured Products, in general, are pre-packaged investments that are based on and linked to some underlying benchmarks, such as equity markets, derivatives, interest rates, commodities, corporate debts, and foreign exchange markets. There is no uniform structure to these investment packages, since they are tailored to meet specific financial objectives. We introduce some typical categories in structured products.

(1) Leveraged Investing

Leverage investing is a technique that invests with borrowed money. The common ones are loans and options/warrants.

- Loans
The mechanism of such leveraged investments can be shown in a simple example, as follows: an investor provides capital of $25M, and a bank provides a loan of $75M. The total amount of $100M is invested in a portfolio of hedge funds over a certain time horizon, 5 years for example. The investment pays interest to the lender, and return the principal at maturity. This principal is paid out of the liquidation value of the investment. After all the obligations are paid off, the investor who initially provided the $25M gets to keep all what is left in the fund.

The structure of such loans can be a bit more complicated, simply because in the simple example mentioned above, the bank takes too much risk and is not usually willing to give the loan without further covenants. The most usual one is one of simple protection, and the bank would chose a trigger loss level, say $5M in the example above, so that when the value of fund falls below the value of the trigger, the fund will be liquidated; in this way, the investor will lose the trigger amount, plus expenses, but the bank does not incur into any loss, provided the fund can be liquidated fast enough. When the fund performs well, the investor only needs to return the principal plus interest and keep the rest as profit. Effectively, the investor is essentially buying a call option from the bank, with the hedge fund as underlying asset, loan amount plus interest as strike price, and option premium as the investor’s initial capital outlay $25M. Because of this, options, which we describe below, are used more often, as they also provide additional advantages for the issuing bank.

- Options/Warrants

Investors can purchase a call option issued as warrants, which provides positive, leveraged returns if underlying asset value increases. The
underlying asset can be virtually anything, and in our case, it is the hedge fund portfolio.

For the loan example in the previous section, instead of the bank providing the $75M, collect the $25M from the investor, and then invest the total $100M in the hedge fund, the bank will prefer to issue a call option to the investor, with a value of $25M, and offer the return of a $100M investment in a fund of hedge funds above the $75M financing fee, which acts as the strike price of the call option (for the sake of clarity, we are ignoring interest rates and fees in this example). This way, the cash flow to the investor are the same, but because an option is issued, the bank does not need to invest the full $100M in the fund; banks hedge their options by taking a position in the underlying instrument equal to the delta of the option, which is less than one (usually close to 0.7) which leads to a considerable savings in bank capital.

We use a small example to explain leveraged return, compared with the return if investing in the underlying directly.

A hedge fund share is now worth $20. An investor buys a call warrant linked to this hedge fund portfolio. This call warrant gives the investor the right to purchase fund shares for $20 from the hedge fund. Suppose now the warrant is worth $0.5. The next day, share value of this hedge fund goes up to $21, and the warrant is worth $1.

Now the investor can sell the warrant for $1. The investment and return schedule is shown in [Table 4].

<table>
<thead>
<tr>
<th>Item</th>
<th>Fund Shares</th>
<th>Call warrant</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial Price</td>
<td>20</td>
<td>0.5</td>
</tr>
<tr>
<td>Sell Price</td>
<td>21</td>
<td>1</td>
</tr>
<tr>
<td>Investment Profit</td>
<td>1</td>
<td>0.5</td>
</tr>
<tr>
<td>Return</td>
<td>5%</td>
<td>100%</td>
</tr>
</tbody>
</table>

Notice that absolute value of return of the warrant is less than that of the direct share investment. This is always true because option value changes less than the underlying value change, and the ratio is the widely know delta of the option. Delta of a option measures the change of the option value with respect to a change of the underlying security, defined as $\frac{\partial O}{\partial S}$, where $O$ is the option value, and $S$ is the value of the underlying security. However, more importantly, the percentage return in the warrant is much higher, 100% compared with only 5% if investing in fund share directly. The 100% return is referred to leverage return, because it only invest in a small portion of the underlying asset value, but still participate fully in the asset value change. This disproportionate participation is exactly the leverage effect, and the value of the option is very sensitive to change in the value of the underlying.

The above scheme looks extremely attractive, but issues exist in leverage investment such as call warrants.
Firstly, leverage always works both ways. Because of the small capital outlay, in this case only $0.5, the return in such investments is very volatile. Imagine now the fund share values drop to $19 instead of rising to $21. A different investment and return schedule is shown in Table 5.

**Table 5.**

<table>
<thead>
<tr>
<th>Item</th>
<th>Fund Shares</th>
<th>Call warrant</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial Price</td>
<td>20</td>
<td>0.5</td>
</tr>
<tr>
<td>Sell Price</td>
<td>19</td>
<td>0.01</td>
</tr>
<tr>
<td>Investment Profit</td>
<td>−1</td>
<td>−0.5</td>
</tr>
<tr>
<td>Return</td>
<td>−5%</td>
<td>−100%</td>
</tr>
</tbody>
</table>

In this case, the call warrant is out-of-the-money and has minimal value, and here we suppose it is 0.01. The loss of investing in the underlying asset is only 5%, while the investor would lose the entire premium paid for the warrant, which translates to a 100% loss.

The second problem is related to hedging the long position in such a call warrant. The usual hedging strategy of a call option is short delta units of the underlying asset. Deltas of options on listed stocks are calculated applying Black-Scholes model, using extensive information on the stock price movements, return behavior, and volatility processes. However, hedge fund shares are not listed, and illiquidity of hedge fund shares give rise to higher implied volatilities. These factors make hedging such a call warrant on hedge fund shares very difficult.

- **CPPI Options**

  Asset allocation between the two parts of the portfolio is done through dynamic trading strategies. One popular strategy is CPPI, which stands for Constant Proportion Portfolio Insurance. This strategy aims to deliver the same as a traditional option, with two major differences:

  - Interest rate hedging. In a traditional option, the issuer of the option is in charge of hedging the interest rate risk inherent in their hedging strategy for the underlying. In a CPPI option, the interest rate risk will be passed over to the investor, or purchaser, of the option, as we will see below.

  - Hedging of the underlying fund returns. In a traditional option, the issuer does not fully invest the option notional in the underlying, but the delta of the option, as dictated by the Black-Scholes strategy. As the performance of the underlying increases or decreases, the hedging strategy results in buy or sell orders of the underlying, to maintain the exposure equal to the delta, which will vary over time. This poses a problem for the issuer, as hedge funds may not have the liquidity required to adjust the hedge continuously; in other words, the hedging strategy may will typically be subject to updates every month, quarter or year, depending on the liquidity of the underlying hedge funds. In a CPPI option, this hedge risk will also be passed to the investor.
In order to explain how a CPPI option will work, consider the following example of a secured loan.

Suppose an investor has $25M and borrows $75M from the bank, to invest in a hedge fund. According to the loan conditions, the first losses will be incurred by the investor; the bank will start to lose when the underlying loses 25% (obviously, these numbers will have to be adjusted slightly to take into account fees and interest, which we will ignore in this example). The bank, to protect itself from losses in their loan, imposes a clause that will be triggered if the performance of underlying portfolio declines by, say, 5%; in this case, the bank will proceed to a partial liquidation of the assets, and the proceeds of the liquidation are used to decrease the credit issued to the investor. If the performance drops another 5%, a further liquidation takes place, and so on. This is done in such a way that, if the performance of the underlying drops by 25% (which is the threshold at which the bank will start to lose on their loan) the bank will have redeemed 75% of the assets, which is the totality of the loan; in this manner, if the underlying fund drops by 25%, the investor would have lost everything, the investment is liquidated, and the bank will lose nothing. Similarly, when the fund is performing well and over above certain trigger levels, the bank will increase the loaned amount and invest it on behalf of the investor. This has the effect of correcting for possible earlier liquidations due to poor performance, but the ultimate effect is maintain the leverage of the loan at constant levels, hence the term: Constant Proportion Portfolio Investment (CPPI).

This investment objective is achieved in a CPPI strategy by issuing an option, as we presented earlier, but with a strike price and underlying notional which is revised every time interest rates fluctuate and every time the underlying investment drops (or increases) beyond a trigger. The strike price is changed to ensure that the delta of the option remains constant; this essentially implies that the bank will minimize the trading activity arising from the hedge and variations in the delta, effectively passing the risk over to the investor.

The effect of CPPI strategies on investors is that their value will decrease very quickly when the fund exhibits losses, but can be extremely profitable if the underlying has good performance. These strategies were very popular in the early decade of 2000, due to the low interest rate environment and high hedge fund returns.

(2) Guaranteed Notes

• Definition of a guarantee

A guaranteed note is an investment strategy where the issuer guarantees the investor that, in case the underlying fund exhibits losses, the investor is guaranteed to get the original value of the investment back to the investor (without interest, though). A guarantee is a structure that combines equity or options on equity, and a fixed income instrument on the equity, in order to protect the principal invested. Guaranteed notes are, strictly speaking, not leveraged structures; however, they obtain their guarantee through the use of leverage, as we will see below.
Guaranteed notes are issued for two main reasons: regulatory environment and risk perceptions. About the first reason, a guaranteed note acts like a bond, which is an allowed investment strategy for most institutions, whereas hedge fund investments are often restricted investment classes. About the second, guaranteed notes will be purchased by investors who consider their cost to be advantageous when compared with the risk the eliminate, and they will be underwritten or issued by institutions who have the exact opposite view on their cost.

Some guarantees are provided by well-rated banks, while others are not. In this regard, it is worthwhile mentioning the case of Portus Asset Management, a Canadian Investment firm that issued guaranteed notes to investors; unlike most guaranteed notes, which are guaranteed by a well rated bank, and not the management firm that creates them, this one was guaranteed by Portus, an unrated company. When Portus defaulted, investors did not get their original investment, simply because the issuer of the guarantee, Portus, was bankrupt. In other words, a guarantee eliminates market risk, but does not eliminate credit risk.

- Anatomy of a guarantee

A portfolio that constructs a guarantee contains two parts: zero coupon bond, and shares of underlying hedge fund, indices, or other assets. Intuitively, the breakdown is exhibited in [Figure 5].

![Figure 5. Anatomy of a guarantee](image)

The zero coupon bond or note is used to guarantee principal in the future. How much amount to invest in this part is determined by interest rate, cash flows, and maturity date, and is set aside to ensure that it will grow in a risk-fashion to guaranteed the return of the original investment at expiration. The part invested in hedge funds is used to obtain exposure in the hedge fund, and it is the aggressive part of portfolio, seeking for higher returns. The exposure is achieved by using using a wide varieties of leveraged financial strategies, such as the ones we mentioned earlier in
this article (non-recourse loan, call option, CPPI option) or the one we mention below: the CFO's.

(3) CFO - Collateralized Fund obligations

- Definition and structure
  A Collateralized Fund Obligation is an interesting structured finance product which can be classified as a type of collateralized debt obligation, commonly known as a CDO. A CDO issues securities or notes backed on a diversified pool of loans, bond, receivables, future flows, or any type of cash flow stream that can be identified and isolated.

  The capital structure of a CFO is similar to traditional CDOs, meaning that investors are offered a spectrum of rated debt securities and equity interest. The difference between CFO and CDO is the collateral that is put forward. In CFO, any managed fund can be the source of collateral. More specifically, the target collateral tends to be hedge funds, such as relative value hedge funds, event-driven hedge funds or commodity trading advisors (CTAs), along with funds that finance the needs of growing companies, such as private equity and mezzanine funds. On the contrary, collaterals of CDOs are mainly loans, structured securities, and debt. For example, the collateral can be mortgage-backed securities, corporate loans, or emerging-market sovereign debt.

- Anatomy of a CFO
  The schematic prototype of a CFO is shown in [Figure 6].

  First of all, a firm set up a Special Purpose Entity (SPE) by transferring some assets to it, and in this case the assets are a pool of hedge funds. This Special Purpose Entity is a separate, independent legal entity, and its profit and loss does not show up in a firm's financial reports. The entity is used to finance special projects, without affecting the entire firm as a whole.

  The SPE is then divided into tranches by different credit ratings, and this provides different levels of yield and risk for investors’ needs. The yields are often quoted as benchmark, such as LIBOR, plus a certain credit spread. The most senior tranche is usually rated AAA and is credit-enhanced due to the subordination of lower tranches. For example, its yield might be quoted as “LIBOR + 0.75%”. This means that the lowest tranche, which is typically the equity tranche, absorbs losses first. When the equity tranche is exhausted, the next lowest tranche begins absorbing losses. In our exhibit, the SPE is divided into four tranches: AAA-rated, AA-rated, A-rated, and Equity tranche.

  With such a SPE set up, there are lenders (issuers) that earn a spread over interest rate. Then there are equity investors, earning total return of the fund minus the financing costs paid to lenders.

- Benefit
  Both investors and issuers can find CFO securitization attractive. For Investors, a triple-A-rated CFO will have a similar yield to that of a triple-A collateralized debt obligation plus a premium. Because of the CFO structure, investors gain exposure to a diversified collection of hedge funds through a fund of funds manager. For issuers, packaging a pool of assets into a SPE and thus obtaining a high credit rating is an ideal way
of raising funds from lenders. This process of constructing a structure product such as CFO is called securitization, and is primarily used for raising funds for an otherwise relatively illiquid product.

- Problems and valuation difficulties

In a collateralized fund obligation (CFO), different hedge funds are pooled in a fund that is in turn securitized. However, it is very difficult to predict the credit solvency of each tranche of the securitization, and therefore give rise to difficulties in pricing CFOs.

More specifically, the credit rating of a CFO tranche depends directly on the mark-to-market value of the pool of hedge funds. The fund of funds manager focuses efforts on actively managing the fund to maximize total return while restraining price volatility within the guidelines of the CFO structure. The huge number of hedge fund investment strategies and the constant change of sell/buy positions among strategies entail that returns of hedge funds are not expected to follow conventional behaviors like traditional assets, such as equities, bonds, and ETFs. Instead they are supposed to be more volatile, unpredictable and should present more general distributions than the usual multivariate Gaussian approach. Therefore the probability of default of the different tranches will be also very volatile and unpredictable.
In addition, the lack of transparency characteristic of hedge funds impedes the acquisition of high-frequency historical data. However, investors and researchers need to calibrate rather complex valuation models on large sample of historic data. This means they would have to compromise to the scarce monthly data in developing and validating the pricing models.

1.6. Assessing the risks. Hedge fund investments share qualitative properties which are quite different from other, more traditional investments.

(1) Illiquidity

Liquidity in general refers to the ability to quickly covert securities or other assets into cash. In hedge funds, specifically, it refers to the length of notice periods that investors need to inform the fund managers before the fund share can be redeemed. This time can vary from a few weeks to a few months.

For example, in our snow fund we would have difficulty turning the swaps into their cash in mid-year, since we would have to negotiate with our counter-parties to cancel our swap contracts with them. This will invariably lead to heavy losses.

Therefore, hedge funds typically have liquidity provisions that will only allow investors to redeem their shares at certain points in time, often with an additional advance notice period required. They may also have “lock-ups”, time from the initial investment that new investors have to wait before their fund shares can be redeemed, according to the usual liquidity provisions of the hedge fund. For example, a fund with a one-year lock and monthly liquidity is a fund that the investor can redeem their shares only after the first year anniversary of their investment and monthly afterwards.

(2) Valuation difficulties

Hedge funds give rise to valuation issues, because sometimes the value of the fund’s assets are hard to value.

In our show fund example, valuing the swap is difficult, not just because of its expected cash flows are uncertain, but also because there is no accepted accounting methodology to value the swaps. Most often, the valuation methodology is established in the Offering Memorandum, and is key to ensure investor equality, as they subscribe and redeem shares.

Evaluating a hedge fund using specific methodologies is obviously more difficult and less transparent than evaluating a public company. In the case of public companies, share prices are simply their market price quoted in the stock exchange. They are determined by supply/demand relationship for the stock, which changes based on investors’ perspective on the value and earning potential of the underlying company.

(3) Capacity restrictions

A fund typically has a limit as to what it can do with its money. For example, in our snow fund example we may have the ability to do 100 swaps at a value of $2M each; this means that if we attract investments much in excess of $200M, the returns will not be as good as if we limit the asset base to $200M. Because of this, many hedge funds managers put a limit to the volume of assets they accept into the fund. Investors must be careful, as some other managers may decide increase these maximums
beyond the natural limits inferred from their investment opportunity, since
the management and performance fee is asset based and will reward the
manager independent of the percentage return of the fund.

(4) Leverage and speculative investment

Hedge funds together with other speculative investments differ from
the old-fashioned investments that do not rely on borrowing capital: hedge
funds are typically highly leveraged. They aggressively take large positions
with only small amount of equity capital, hoping for large returns with
even very small changes in favor of their positions. The high leverage gives
rise to leverage risk, which means a small unfavorable change can trigger
enormous losses as well, especially if positions are run unhedged.

There are also other risks involved apart from the main ones mentioned above.
For example, most hedge funds trade on secretive strategies, causing less trans-
parency for investors. The fact that hedge funds have full authority over their own
trading styles and most actions taken are regulated make the funds vulnerable to
attacks and crashes in the financial markets, such as the crisis 2008.

1.7. Hedge Fund Data. Hedge funds, as private investment partnerships,
are under no obligation to disclose investment information to anyone, and the dis-
closure to their own investors is often limited to the NAV, monthly share price and
little else. Even their investment strategies are often kept within close guard inside
the management company. As a result, one of the few pieces of information that
one can find about a particular hedge fund, not being an investor, is the series
of monthly returns. Moreover, publications of hedge funds are considered illegal
marketing materials in some jurisdictions such as the U.S, and is therefore heavily
restricted to the extend that they cannot be published by the management com-
panies in public forums, websites, etc. Marketing of hedge funds is restricted to
certain individuals and cannot be brought freely to the public.

However, as a miracle of the law, publication of such data by third parties is
allowed. Commercial databases are accessible to qualified users who paid a fee to
use the data sets. These commercial databases exist and we name just a few:

- Hedge Fund Research Inc.
- Hedgewfund.net
- Eureka hedge
- Tass

To understand the informational content of such databases, one needs to un-
derstand the following well-known syndromes:

- Backfill bias. Hedge funds are under no obligation to report to a database.
  Because of that, they choose when they report. They often do so when
  they have good performance they want to boast about, usually in times
  when they are looking to gather more investors.
- Survivorship bias. Hedge funds which collapse tend to disappear from the
databases.

2. The statistics of fund returns

As we mentioned earlier, hedge fund share price contains most of the informa-
tion of the fund, reflecting the value of investment in the fund, but if we want to
get a quantitative idea of the performance of the fund, we have to introduce new
HEDGE FUNDS

concepts. The first is the *return*. Intuitively, the *return* is a mathematical term that embodies the growth characteristics of the fund’s share price over time. In its simplest characterization, the monthly return is defined as follows:

**Definition 2.1.** For a given month \(k\), the fund’s return \(r_k\) in that month is calculated as

\[
r_k = \frac{S_k - S_{k-1}}{S_{k-1}},
\]

where \(S_j\) denotes the fund’s share price for month \(j\).

**Definition 2.2.** For a fund with monthly returns given by \(r_k\), from \(k = 0, \ldots, m\), the arithmetic mean return is given by

\[
R_{amr} = \frac{\sum_{k=1}^{m} r_k}{m}.
\]

In Example 2.3 and Example 2.4, it shows that when simple return measure might not be in line with the real investment performance. As intuitive as this definition may be, it has serious limitations.

**Example 2.3.** Imagine I set up my own fund with my own $1 as the only initial investment, with one share, in January 1, and imagine also that every month until December I manage to double the value of my investment, *without the inflow of any other assets*; in other words, the share value at the end of November is \(2^{10} = 1,024\). The monthly return is equal to 100%. A wealthy friend of mine, impressed by my rate of return, decides to invest $1,000,000 = $1M on December 1st. The total asset value at that time is $1,001,024. At that time, I lose half of the assets of the fund, and my return for December is -50%. My average monthly return is

\[
\frac{11}{12} 100\% + \frac{-50}{12} = 87.5\%.
\]

This number shows a significant positive return, when in fact the investment ended up with a net investment loss of $499,489.00 during the year.

**Example 2.4.** Consider a fund that has $1 in assets invested in securities, with the following results; the first month, the value of the securities double, to a total of $2; the next month, the value of the securities is cut in half, back to $1; the month after they double again, to be cut in half again the month after that; and so on and so forth. The monthly returns of the fund will then be +100%, -50%, +100%, -50%, etc. Nothing wrong with these numbers, but if one decides to calculate their average, one will find that the *average return* of this fund is +25% per month; a little strange for a fund that makes no money.

Later, when we tackle the problem of doing statistics properly on fund returns, we will see that this is an issue that we have to live with. However, there is an alternative definition of returns that is sometimes used, and referred to as *log-returns*.

**Definition 2.5.** For a given month \(k\), the fund’s log-return \(r_k^{\text{log}}\) in that month is calculated as

\[
\begin{align*}
(2.1) & \quad r_k^{\text{log}} = \log \frac{S_k}{S_{k-1}}, \\
(2.2) & \quad = \log(1 + r_k).
\end{align*}
\]

where \(S_j\) denotes the fund’s share price for month \(j\).
Note that, by Taylor’s theorem,
\[ r_k = r_k^{\log} + O(r_k^2), \]
which means that, for return time series which do not have large swings (i.e., volatility), both numbers are actually quite close.

A question we may ask is: why don’t we simply use the log-returns and ignore the simple return introduced above? Example 2.6 provides the answer.

**Example 2.6.** Assume we are invested equal amounts in two assets A and B. Say that asset A doubles in value, while asset B drops in value 50%. One can easily calculate that the portfolio will make a 25% return measured with standard returns, which happens to be exactly the average of the individual returns of A and B respectively. However, if we use log-returns, the average of the log-returns of A and B is 0, but the log-return of the portfolio is strictly positive (22.31%).

This example shows that although log-returns may be more appropriate sometimes when analyzing individual asset performance, simple returns are still more useful in portfolio theory, when we look at combinations of individual assets.

Apart from simple return and log return, some other return measures are developed, in trying to capture the real return characteristics over time. As we will see, they are not perfect either.

**Definition 2.7.** For a fund with monthly returns given by \( r_k \), from \( k = 1, \ldots, m \), the time-weighted return is given by
\[ R_{\text{tw}} = \prod_{k=1}^{m} (1 + r_k) - 1. \]

From the definition, we can see that the time-weighted return measures compound growth in a portfolio. It is also called geometric mean return, as opposed to simple arithmetic returns.

When we want to account for all the cash flow generated by the investment over time, we may prefer the following definition, and solve the present-value problem.

**Definition 2.8.** For a fund with monthly cash flows given by \( N_k \), from \( k = 0, \ldots, m \), the Internal Rate of Return (IRR) is defined implicitly to be the rate \( R_{\text{irr}} \) that equates the ending value of investment to the future value of all the investment cash flows.
\[ N_m = \sum_{k=0}^{m} (1 + R_{\text{irr}})^{m-k} N_k. \]

Notice that IRR defined for a fund has great similarity to the Yield To Maturity (YTM) for bonds. For a bond with market value \( P_0 \), and cash flows at each period \( CF_k \), the YTM is defined as the rate \( r \) that equates the market value to the present value of the cash flows:
\[ P_0 = \sum_{k=0}^{m} \frac{CF_k}{(1+r)^k}. \]

As a matter of fact, the concept of Internal Rate of Return (IRR) applies widely to all kinds of investments, and is also used in a firm’s capital budgeting process. It intuitively represents the rate of return implied by the cash flows of
an investment, and it is considered to be a comparable, quick way of evaluating investment performance.

The choice between different return definitions is entirely a matter of the type of portfolio performance information one is looking for. We need to understand that return numbers may not reflect the real investment returns closely, and if we use the numbers alone without a reference to the actual value of the assets invested, the conclusion might be misleading.

To report the cumulative performance of a fund over time, a measure called "Value Added Monthly Index" (VAMI) is often used, which tracks the monthly growth of a hypothetical $1,000 investment. Its value $V_{k_0,k}$ at a given month $k$, for an investment start date $k_0$, is given by

$$V_{k_0,k} = 1,000 \cdot \prod_{m=k_0+1}^{k} (1 + r_m)$$

**Definition 2.9.** For a fund with returns $r_k$, the mean, or expected value of the return series is defined as

$$\mu = \mathbb{E}[r]$$

The unbiased, efficient, and consistent estimator of the mean is the sample average, which is

$$\bar{r} = \frac{\sum_{k=1}^{m} (r_k)}{m}$$

The formula is identical to arithmetic mean that we previously defined. The mean return measures the central tendency of a series, or in other words, the level of return that can be expected on average.

**Definition 2.10.** For a return series $r_k$, standard deviation is a statistical term defined as

$$\sigma = \mathbb{E}[(r - \mu)^2]$$

Its unbiased sample estimator is

$$S = \sqrt{\frac{1}{n-1} \sum_{i=1}^{n} (r_i - \bar{r})^2}$$

**Definition 2.11.** For a return series $r_k$, covariance is a statistical term describing how much two variables move together. For return series $R_x$ and $R_y$, the covariance between the two series is defined as

$$\text{COV}(R_x, R_y) = \mathbb{E}[(R_x - \mu_x)(R_y - \mu_y)]$$

Another concept, correlation, is a standardized measure of covariance, defined as

$$\rho(R_x, R_y) = \frac{\text{COV}(R_x, R_y)}{\sigma_x \sigma_y}$$

### 3. Types of Hedge Funds

Perhaps our snow fund example earlier in the paper embodies the main characteristics of what a hedge fund is. But most hedge funds are not like that, they are based on some of the more traditional investment vehicles, such as stocks, bonds, convertible securities, futures, etc. Because of that, they are usually classified into groups of more or less homogeneous characteristics, usually making a reference
to the type of securities they use to obtain performance. In this section we will introduce the general classification and then describe some main hedge fund styles.

3.1. General Classification.

3.1.1. Equity Hedge. This refers to hedge fund strategies that set up both long and short positions in equity and equity derivative securities. Long and short positions need to chosen carefully so that the long position provides return to the portfolio and short position acts as a hedge against stock market decline.

Ideally, it should work in the following way: in a rising market, the long-position should increase in value faster than the market returns, and the short position incur a loss less rapidly. In this case, the return from the long position is reduced by the loss in the short position. In a declining market, the short position holdings should decrease in value faster than market, generating a big return for the portfolio. The long position should decrease in value less rapidly than the market, so the loss is limited. In this case, the long position is hedged by the short position, and that is why fund managers are willing to accept return reduction in rising market for this protection when market declines.

In this general category, investment decisions are reached by applying two main techniques: identifying economic trends, and fundamental analysis. Fund managers first predict the effect of macro-economic trends on the market, and then identify industries that are attractive for long positions. Within these industries, the fund manager then seeks for well-positioned companies in the industry, usually through quantitative analysis of the company’s financial status and earning prospects.

Equity hedge strategies include strategies such as Equity Long/Short, Quantitative direction, and Fundamental Value. We will discuss Equity Long/Short in more details.

3.1.2. Event Driven. Event driven strategies are hedge fund strategies that exploit pricing inefficiencies caused by a wide range of corporate transactions, such as mergers, restructuring, and financial distress. The fund managers develop strategies to take advantage of special situations in corporations that change the value of its underlying assets.

Event Driven strategies include some major strategies such as Distressed Securities, Merger Arbitrage, and Regulation D.

3.1.3. Relative Value Arbitrage. Relative Value Arbitrage strategies seek return from "spreads". That is, they derive return from price relationships between two related securities. The arbitrageurs seek distorted relationships between securities by quantitative analysis and valuation, take long and short positions in them, and then realized a profit when the relationship returns to normal.

Relative Value strategies include but not limited to Fixed-Income Arbitrage, Convertible Arbitrage, Asset Backed Arbitrage, and statistical arbitrage.

3.1.4. Global Macro. Global Macro strategies often include Active Trading, Commodity Trading, Currency Trading, and Multi-Strategy. We will introduce this category in detail later.

In the following sections, we will discuss in detail some main strategies from each of the general categories. We will introduce their hedge approaches, source of returns, and finally risk and risk controls.

3.2. Equity Hedge - Equity long/short. These funds obtain performance for their investors by buying and selling stocks. The net exposure of the long and
short positions varies by funds’ investment schema and managers’ trading style. At first sight, they exhibit a fundamental difference with mutual funds: their ability to short sell stock.

3.2.1. Short Selling. Short selling stock is a common practice among equity investors. An investor short sells stock when she believes that stock price is going to decrease. When the price does drop, she can purchase the stock at the lower price and make a profit from the difference between the short proceeds and the lower repurchase price. Mathematically, we can think of short selling stock easily by simply assuming one can buy a negative amount of stocks. However, we need to examine further the actual investment process of short selling stock, as it will have important consequences in trying to understand this particular trading style.

Stock borrowing, or short selling, is usually done through a broker. Hedge funds hire typically one or more prime brokers, who are mainly responsible for trading securities, and often they also maintain custody of the assets. When a manager wants to short sell a stock, she makes arrangements through the prime broker to borrow stock from an existing shareholder. Primer brokers typically have ample inventories of stocks from their clients that are ready to be lent. This requires prior approval of the owner of the stock, and the stock loan can be recalled anytime by the original owner. After the broker has finalized the stock lending process, typically in a matter of a few minutes (except for so-called hot securities, when it can take hours or days), the stock can be sold and the proceeds of the sale go into the fund account. The fund, at that time, has a credit equal to the cash proceeds of the sale and a debit equal to the shares they borrowed.

A short position is usually unwound when the borrower decides to buy the shares and give them back to the original owner. However, we should be aware of the potential risks in shorting stocks. If the stock price increases significantly, some sellers are forced to cover the position in case of a margin call (a monetary deposit required as collateral by the broker in situations where losses are unrealized but large). Also, the original owner of stock may decide to sell the stock and make a profit. Then the borrower is forced to buy the stock to return it to the original owner. Buying shares further increase the stock price and trigger even more position covering. This event is referred to as a short squeeze, and it usually has very negative implications for the borrower. A recent example occurred in 2008, with Volkswagen stock (VW), and happened as follows:

On Friday October 24, 2008, VW was owned largely by Porsche, who owned 42.6% of the stock, and the state of Lower Saxony, who owned 20.1% of the stock. Because of the poor prospects of the automobile industry at the time, Volkswagen was the most shorted stock in Germany’s benchmark DAX Index with about 12.9% of Volkswagen’s stock on loan, mostly for short sales. In an attempt to take over VW, Porsche announced on Sunday October 26 that it effectively controlled 74.1 percent of VW stockholdings, 42.6 percent of the carmaker’s common stock, and 31.5 percent in cash-settled options. This left less than 6 percent of Volkswagen stock float, or stock not held by majority owners or insiders. Counter-parties to the VW options started long the stock to hedge their positions, leaving fewer stocks available to buy back. Stock price rose, owners of the stocks wanted to sell, and stock recalls were issued to the short sellers. This triggered demand from short sellers, further contributing to the rise of the VW stock price, triggering more recalls and more demand from all short sellers. This chain effect resulted in Volkswagen’s
skyrocketing stock price. Volkswagen shares soared 286 percent in the week ending Tuesday, closing at a record high of 945 euros (about $1,198) on Tuesday. The short squeeze in Volkswagen, pushed Volkswagen’s valuation as high as $370 billion, above the market cap of Exxon Mobil (NYSE: XOM), normally the largest publicly traded company in the world.

3.2.2. Short sell proceeds and leverage. Some of the delicate nuances of equity long/short managers is using the proceeds from short-selling to buy other stocks, and the possibility for leverage. To better understand this, we examine the following situation.

Example 3.1. A manager has $1,000 to invest. There are two stocks that she is analyzing: Superbco trades for $100 and she thinks the stock will outperform, and Horrendco, also trading at $100, she thinks will underperform. So she short sells 9 shares of Horrendco and uses the proceeds to buy 9 shares of Superbco. Let’s assume the manager deposits the $1000 cash at risk-free interest rates, 2%. The the stock borrowing fee is the same 2% (payable to the stock lender) plus an additional 1% payable to the broker, which is a total cost of 3% for the short selling.

One year later, the manager’s thesis proves to be correct, and indeed Superbco’s price increases to $110, whereas Horrendco increases to only $105. For consistency, let’s assume the overall market rose by 7.5%. Since the manager was long Superbco and short Horrendco, the manager made $90 and lost $45 on the stock pair trade, had to pay a borrowing fee of $27, and made 2% interest on $1,000, leading to a net gain of $38.

The cash flows of this trade is summarized in Table 6

<table>
<thead>
<tr>
<th>Item</th>
<th>Before the trade</th>
<th>After trade</th>
<th>A year later</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cash</td>
<td>1,000</td>
<td>1,000</td>
<td>1,020</td>
</tr>
<tr>
<td>9 Superbco</td>
<td>0</td>
<td>900</td>
<td>990</td>
</tr>
<tr>
<td>9 Horrendco</td>
<td>0</td>
<td>−900</td>
<td>−945</td>
</tr>
<tr>
<td>Borrowing fee</td>
<td>0</td>
<td>0</td>
<td>−27</td>
</tr>
<tr>
<td>Total</td>
<td>1,000</td>
<td>1,000</td>
<td>1,038</td>
</tr>
</tbody>
</table>

The manager obtained a return of 3.8% on this trade. At first sight, this does not seem interesting; return of this trade is well below market performance 7.5%. However, if the market had declined, instead of a rise of 7.5%, this trade would have apparent advantage. Since this trading strategy is based on the relationship between the stocks, the return will hold provided that the spread between the two stocks is constant. Consider the following variation:

Example 3.2. A manager invests $1,000 as in Example 3.1. Now Superbco decreases to 95 and Horrendco decreases to 90. For consistency, we assume the market performance decreased by 7.5%.

In this case, cash flows of this trade is summarized in Table 7
Table 7.

<table>
<thead>
<tr>
<th>Item</th>
<th>Before the trade</th>
<th>After trade</th>
<th>A year later</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cash</td>
<td>1,000</td>
<td>1,000</td>
<td>1,020</td>
</tr>
<tr>
<td>9 Superbco</td>
<td>0</td>
<td>900</td>
<td>855</td>
</tr>
<tr>
<td>9 Horrendco</td>
<td>0</td>
<td>−900</td>
<td>−810</td>
</tr>
<tr>
<td>Borrowing fee</td>
<td>0</td>
<td>0</td>
<td>−27</td>
</tr>
<tr>
<td>Total</td>
<td>1,000</td>
<td>1,000</td>
<td>1,038</td>
</tr>
</tbody>
</table>

The only difference in Table 7 from Table 6 is the share value of Superbco and Horrendco, 855 and -810 compared with 990 and -945. However difference in the share values remain the same $45. We see that the return is the same, 3.8%, and it now beats the overall equity market by 11.3%. More importantly, this strategy has positive returns no matter the market goes up or down: return is generated as long as manager’s prediction of the stock relationship is correct: Superbco is better than Horrendco. This is an important feature of equity long-short managers, and their return can be independent of market directions.

However, to understand how returns can be attractive compared with interest rates, their trade decisions involve another important aspect: their ability to manage cash effectively. To get us in the right mind frame, observe that it seems that the $1,000 that we started with are unnecessary. We could, in principle, short sell Horrendco and use the $900 short proceeds to buy Superbco; in this way, the profit of $38 would represent a return on zero initial capital, which is a return of ∞%. Or, equivalently, can we buy more than just 9 stocks, but 90, or 900, or 900,000, and short the same amount? However, if our strategy does not work, it will trigger bankruptcy, and our prime broker will be, at best, unhappy, as they will have to absorb the losses. This leads to the fundamental question; how much money do we need to run this strategy? or, equivalently, how many stocks can we buy or sell with existing capital?

This question is deeper than it may seem. In fact, investment banking, which was largely unregulated in the eighties and nineties, did allow for the prime broker to use their own discretion to determine required capital from their clients, who provide the funds, to perform different transactions. This permitted funds such as Long Term Capital Management to construct extremely large portfolios with relatively small capital bases. At the end of the day, the capital required will have to make sure that under most scenarios the fund will be able to meet its liabilities in the case of portfolio losses.

To make a long, complex story short and easy, let’s assume a very simple collateral requirement by the prime broker.

- Long positions require 50% of the purchase price as collateral.
- Short positions require 80% of the short-selling price as collateral.

If, as the previous examples, we want to buy 9 stocks valued at $100 and short sell another 9 stocks valued at $100, collateral requirements will impose a balance of at least $450+$720=$1170. Since we will obtain $900 as short selling proceeds, the
least amount of cash needed to run this strategy is $270. Therefore, the following example is completely feasible.

**Example 3.3.** A manager invests $270 as in Example 3.1.

In this case, the cash flows are shown in Table 8.

<table>
<thead>
<tr>
<th>Item</th>
<th>Before the trade</th>
<th>After trade</th>
<th>A year later</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cash</td>
<td>270</td>
<td>270</td>
<td>275.40</td>
</tr>
<tr>
<td>9 Superbco</td>
<td>0</td>
<td>900</td>
<td>990</td>
</tr>
<tr>
<td>9 Horrendco</td>
<td>0</td>
<td>-900</td>
<td>-945</td>
</tr>
<tr>
<td>Borrowing fee</td>
<td>0</td>
<td>0</td>
<td>-27</td>
</tr>
<tr>
<td>Total</td>
<td>270</td>
<td>270</td>
<td>293.4</td>
</tr>
</tbody>
</table>

This is important: our new profit of $23.4, although smaller than the profit in example 6 because we get less interest from the smaller amount of cash we keep, this investment returns 8.66%, well in excess of the 3.8% return in Example 6.

The situation in 8 is very common and is another example of a leveraged —or levered— investment, which we briefly touched upon our previous example of the snow swap. The situation in 6 is also common, and is referred to as an unlevered investment, but they are vastly different. The difference manifested itself in very real, scary form around August 8th, 2007. To attempt to explain what happened, we are going to evolve beyond Example 6 in the following, entirely fictional example, but one that embodies the main elements of what happened in that month.

Imagine not just one, but several funds all sharing the same trades as in Example 6, and one of them is a very large fund owned by an investment bank. It owns lots of shares of Superbco and is short many shares of Horrendco, so many as to be a significant part of the total number of shares issued by both companies. Let’s call this large fund GS Fund. Let’s also assume that all the funds were investing with leverage to the limit, as in Example 8. Because of the liquidity problems that the market was going through in 2007, imagine that the clients of the large fund decide to simply redeem their investment in the fund. When that happens, GS sells some of the stock in Superbco (which causes its price to go down) and covers some of its shorts in Horrendco, causing the price to go up. At that point, the prime broker of all the funds re-do their collateral calculation and they realize that the funds are exceeding their leverage limits: the value of their assets minus liabilities is less than the $270 needed to support their portfolio. When that happens, the prime broker issues a *margin call*: they call the fund and ask them to make a payment to the broker, which is used to refill the collateral back to $270. However, these smaller funds do not have any more cash, as they are leveraged to the limit, and their only recourse is to sell Superbco and cover Horrendco, causing their prices to go further up and down, respectively. This causes more margin calls, and very rapidly we see that all the funds must liquidate their positions, and at the same time, without any regard to the quality of the investments, we see the price of Superbco plummet and the price of Horrendco going through the roof.
3.2.3. Risk and Risk Control.

- Stock picking risk
  Managers of Equity Long/Short strategies rely on the correct pick of stocks for long and positions for return of the portfolio. Therefore, managers need to make correct predictions about company earning potentials and stock price movement based on in-depth fundamental research of the companies. Incorrect predictions will incur a direct loss in the long/short positions.

- Market Risk
  By investing in stock market, the Equity long/short strategy is inevitably affected by the overall financial market and industry performance. For example, a fund manager picks a company with favorable earning potentials within a industry for its long position, but a significant bad news happens in the industry, causing stock prices to drop in the industry as a whole. Then the strategy would fail to work as well, causing the strategies of some funds to fail.

- Price manipulation
  This risk is directly reflected in Example 3.3, where a big player can deliberately induce a short squeeze.

3.3. Equity Hedge - Equity Market Neutral. Equity Market Neutral strategy in general refers to pick stocks and construct portfolio that has zero net exposure, so that the portfolio value is not affected no matter of the market direction. This is a strategy that is in contrast with the directional strategies, and we will analyze it in several aspects.

3.3.1. Alpha and beta of a stock. To understand market neutral strategies well, the knowledge and intuition of alpha, beta of a stock is necessary. $\alpha$ and $\beta$ are the two parameters of Security Characteristic Line (SCL) in Capital Asset Pricing Model (CAPM). The model says that in a well diversified portfolio, the return of an asset depends only on its relationship to the overall market, and can be represented by a linear equation consisting of a constant and a proportional term to the market. More specifically, the model uses the sensitivity of a stock return with respect to the market returns as the main driven force of the stock return, and the equation is as follows:

\begin{equation}
    r_{i,t} = \alpha_i + \beta_i r_{M,t} + \epsilon_{i,t}
\end{equation}

\begin{equation}
    \beta_i = \frac{\text{cov}(r_i, r_M)}{\text{var}(r_M)}
\end{equation}

where $r_{i,t}$ and $r_{M,t}$ are the stock and market return at time $t$, and $\beta$ is calculated as the covariance between the stock and the market divided by the variance of market returns. $\alpha$ is the excess return of the stock, after excluding its return related to the market movements. Fund managers look for stocks with positive $\alpha$’s.

To illustrate this concept in practice, we estimate this equation for Apple Inc. (NASDAQ:AAPL) with respect to the S&P returns, which serves as an approximation for market portfolio. [Figure 7] compares their stock prices over years 2003 to 2009, and [Figure 8,9] display their daily returns.
Now we regress the daily return of Apple Inc. on the S&P daily return, to obtain the $\alpha$ and $\beta$ coefficients. Regression results are shown in Table 9.

Table 9.

<table>
<thead>
<tr>
<th>Coefficient</th>
<th>Estimate</th>
<th>Standard Deviation</th>
<th>T-statistic</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha$</td>
<td>-0.0015</td>
<td>0.00058</td>
<td>-2.74</td>
<td>Significant</td>
</tr>
<tr>
<td>$\beta$</td>
<td>1.081</td>
<td>0.042</td>
<td>25.65</td>
<td>Significant</td>
</tr>
</tbody>
</table>
The regression estimates of both coefficients are significant, and estimate for the $\beta$ coefficient is consistent with the widely acknowledged industry estimate Apple Inc.’s $\beta$ value between 1.02 and 1.10. It is worth noticing that the estimate for $\alpha$ is negative and statistically significant. This means, in the past 6 years, the excess return of Apple Inc. is negative and statistically significant. This means, in the past 6 years, the excess return of Apple Inc. Systematic, after accounting for its risk correlated with the market, is actually negative. This is bad news for Apple Inc. share prices, and the -0.15% per day negative excess return translates to a cumulative -2.9% monthly return. However, if one looks at the recent history of Apple, one can see that it has substantially outperformed the market. Future stock behavior (or $\alpha$) is seldom related to its past performance record, although its $\beta$ are more reliable, or stable through time.

3.3.2. Types of managers. There are basically two types of managers in this broad sectors. One of them trades in stocks only, but can take long or short positions. The objective is to hedge the long positions with the short ones, by holding an equal amount in long and short positions, usually referred to as “dollar neutrality”. Systematic risks are expected to neutralized by these offsetting positions, and returns are obtained independent of market direction. The other one takes equity positions (long or short) and hedges the market exposure with futures, indices or derivatives, including options. These are the managers that are responsible for the term hedge fund, although when the term was coined in the seventies they operated mostly inside the private banking group of the large banks. By the word “Neutral”, the strategy concept is extend to neutral in multiple variables, such as:

- Beta
  By making the portfolio beta-neutral or low beta, the portfolio would be less sensitive to market directions, and thus bear less market systematic risk.
- Industry
Rather than betting on the development direction of a certain industry, the fund manager can invest the same amount in long and short positions in the same industry. This neutrality eliminates the industry risks in the investment. More specifically, this can be done by first rank all the companies in the industry, and then select the top ranked ones for long positions, while short the ones with lowest ranks.

3.3.3. Source of Return. Funds of this type strive for return independent of market conditions. The source of return of Equity Market Neutral hedge is from the relative outperformance of long position with respect to short position. That is, there is a positive return as long as the long performance achieves a bigger return than the short position. Such price relationships between securities are identified by employing sophisticated quantitative models to analyze stock information and predict stock price movements.

3.3.4. Correlation with market. Using HFRX daily return data for Equity Market Neutral strategy, we construct its correlation with S&P 500 returns. [Figure10] shows the correlation graph for bull and bear periods separately. Here the distressed time is defined as from September 2007 onwards, and boom (or bull) time is defined as year 2003 to 2007. From the graph, we can see that in bull market, correlation is mainly between 0 and 0.2, with a smaller peak around 0.3. However, in distressed times, the majority of correlation is negative, with some points around 0.1. It seems like that the correlation graph has shifted left in distressed time, providing very low or even negative correlation with the equity market. The implication of this is major: managers of Equity Market Neutral strategies can provide low correlation in market distress, and therefore provide diversification in investment portfolios, when diversification is hardest to find.

3.3.5. Risk and Risk Controls. Although Market Neutral strategy appear to be able to neutralize primary risks, some other exposures still exist.

- Equity Volatility
  Equity volatility is closely related to equity returns, as is well known and we summarize in [Figure 11], a scatter plot of the S&P monthly return and intra-month volatility between 2004 and 2008, as well as a linear regression fit line.

- Model risk
  As mentioned, in this strategy, fund managers of this type resort to quantitative models in identifying investment opportunities. These models give rise to model risks. This includes failure to pin down the model, misapplication, or low prediction of the model.

3.4. Relative Value - Convertible arbitrage. A convertible security is, in a nutshell, a security that combines the best features of a bond with the best features of a stock. The best feature of a bond is that the promise to pay is backed by the firm’s assets; in other words, in the case of default or bankruptcy, bond holders are first in line to collect their debt. The best feature of stocks is that, in good times, shareholders keep all the profits. A convertible bond allows the holder of the security to exchange the bond for a predetermined number of shares at a pre-specified price. The exchange sometimes happens at the discretion of the holder, sometimes at the discretion of the issuer.
Figure 10. Correlation between Equity Market Neutral strategy and S&P 500 daily return

Figure 11. Equity Hedge Index 2003-2008

Convertible securities have been very successful ways to raise capital for companies with poor credit rating and uncertain future - for which bonds or equity alone
would have been difficult to sell- but with strong future outlook that makes the conversion feature attractive to investors, whose investment would be backed by the firm’s assets. One of the best known examples is MCI Communications, who managed to fuel its tremendous growth from 1978, when it was worth $161 million, to March 1983 when it was worth $2.071 billion. An issue of convertible preferred stock in December 1978 raised $28 million, followed by a second in September 1979 that raised $67.5 million and a third offering in October 1980 raising $49.5 million. The conversion feature allowed the firm to call the bonds provided that the market price of MCI stock exceeded the conversion price by a pre-specified margin of around 25% for 30 consecutive trading days around the call date. As events turned out, MCIs stock price rose enough for it to be able to force conversion on all three issues by November 1981. Therefore, the bonds disappeared from the company’s balance sheet and it allowed it to issue more debt in the same fashion going forward.

The convertible arbitrage strategy uses convertible bonds, but they are usually hedged by the investor; they can be hedged in several ways, normally by shorting the underlying common stock. The quantitative valuation is overlaid with credit and fundamental analysis to reduce the risk further and increase potential returns. Often, emphasis is placed on growth companies with volatile stocks, paying little or no dividend, with stable to improving credits and below investment grade bond ratings. Convertible arbitrageurs construct long portfolios of convertible securities and hedge by selling short the underlying stock of each security. Convertible securities include convertible bonds, convertible preferred stock, and warrants. The price of the convertible declines less rapidly than the underlying stock in a falling equity market and mirror the price of the stock more closely in a rising equity market. Arbitrage opportunities are identified using valuation models that locate a cheap convertible security: the market value of the convertible is less then the expected value, given the market price of the underlying security, interest rates, credit quality, implied volatility, and time expiration or call probability. Cash flows for these strategies are usually given by the bond yield, short interest rebate and a small outflow for short stock dividends. Trending stock price movement provides the opportunity for trading profits while holding the arbitrage position. Since hedge adjustments capitalize on stock price volatility, trading profits can be expected during normal market conditions of stock volatility and stable interest rates.

Example 3.4. Consider the convertible arbitrage strategy example of a bond selling below par, at $80.00: it has a coupon of $4.00, a maturity date in ten years, and a conversion feature of 10 common shares prior to maturity. The current market price per share is $7.00. The client supplies the $80.00 to the investment manager, who purchases the bond, and immediately borrows ten common shares from a financial institution (at a yearly cost of 1% of the current market value of the shares), sells these shares for $70.00, and invests the $70.00 in T-bills, which yield 4% per year.

In order to understand the profitability and risks of such strategy, we are going to introduce the concept of scenario generation. In this setting, we will simply consider a collection of different future scenarios for the prices of bonds and stocks, say a year from now. In more realistic or sophisticated situations, the naive scenario generation analysis we will do here will be replaced by a computer program, probably an excel spreadsheet, that will generate thousands of possible future scenarios,
using probabilistic assumptions, and re-do the same analysis we will do below but in a much more exhaustive manner.

Scenario 1: Values of shares and bonds are unchanged:

Table 10.

<table>
<thead>
<tr>
<th>Item</th>
<th>Today</th>
<th>Scenario A</th>
</tr>
</thead>
<tbody>
<tr>
<td>Convert bond</td>
<td>80</td>
<td>80</td>
</tr>
<tr>
<td>Stock</td>
<td>−70</td>
<td>−70</td>
</tr>
<tr>
<td>T-Bill</td>
<td>70</td>
<td>72.8</td>
</tr>
<tr>
<td>Coupon</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>Short selling fee</td>
<td>0</td>
<td>−3.5</td>
</tr>
<tr>
<td>Total</td>
<td>80</td>
<td>83.3</td>
</tr>
</tbody>
</table>

Scenario set 2: the bearish view. In the next two examples, the share price has dropped to $6.00, and the bond price has dropped to either $73.00 or $70.00, depending on the reason for the drop in share market values. This reflects the fact that correlations between stocks and bonds are positive, so they tend to move in the same direction. The net gain to the client is 7.87% and 4.12% respectively, again after deducting costs and fees.

Table 11.

<table>
<thead>
<tr>
<th>Item</th>
<th>Today</th>
<th>Scenario B</th>
<th>Scenario C</th>
</tr>
</thead>
<tbody>
<tr>
<td>Convert bond</td>
<td>80</td>
<td>73</td>
<td>70</td>
</tr>
<tr>
<td>Stock</td>
<td>−70</td>
<td>−60</td>
<td>−60</td>
</tr>
<tr>
<td>T-Bill</td>
<td>70</td>
<td>72.8</td>
<td>72.8</td>
</tr>
<tr>
<td>Coupon</td>
<td>0</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Short selling fee</td>
<td>0</td>
<td>−3.5</td>
<td>−3.5</td>
</tr>
<tr>
<td>Total</td>
<td>80</td>
<td>86.3</td>
<td>83.3</td>
</tr>
</tbody>
</table>

Scenario set 3: the bullish view. In the following three examples, the share price increased to $8.00, and the bond price increased either to $91.00, $88.00 or $85.00, depending on the expectations of investors, keeping in mind that we have one less year to maturity. The net gain to the client is 5.37% and 1% in the first two examples, with an unlikely net loss of 2.12% in the last example.

The reader with a penchant for risk management will quickly see here the beginning of a scenario analysis of a convertible trade that could easily be turned into a Value-at-Risk calculation. Also, it is interesting to note that a correlation breakdown in the bond-equity relationship will also give rise to much unwanted losses.
Table 12.

<table>
<thead>
<tr>
<th>Item</th>
<th>Today</th>
<th>Scenario D</th>
<th>Scenario E</th>
<th>Scenario F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Convert bond</td>
<td>80</td>
<td>91</td>
<td>88</td>
<td>85</td>
</tr>
<tr>
<td>Stock</td>
<td>−70</td>
<td>−80</td>
<td>−80</td>
<td>−80</td>
</tr>
<tr>
<td>T-Bill</td>
<td>70</td>
<td>72.8</td>
<td>72.8</td>
<td>72.8</td>
</tr>
<tr>
<td>Coupon</td>
<td>0</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Short selling fee</td>
<td>0</td>
<td>−3.5</td>
<td>−3.5</td>
<td>−3.5</td>
</tr>
<tr>
<td>Total</td>
<td>80</td>
<td>84.3</td>
<td>81.3</td>
<td>78.3</td>
</tr>
</tbody>
</table>

We will end the scenario analysis considering two extreme situations: the company goes bankrupt, or the company doubles in value. In the case of the bankrupt company, the bond drops to the recovery value, usually half of the face value of the bond. In the case that the stock doubles in price, we will convert the bond into stock. The financials will then look like table 13 below:

Table 13.

<table>
<thead>
<tr>
<th>Item</th>
<th>Today</th>
<th>Scenario G</th>
<th>Scenario H</th>
</tr>
</thead>
<tbody>
<tr>
<td>Convert bond</td>
<td>80</td>
<td>50</td>
<td>140</td>
</tr>
<tr>
<td>Stock</td>
<td>−70</td>
<td>0</td>
<td>−140</td>
</tr>
<tr>
<td>T-Bill</td>
<td>70</td>
<td>72.8</td>
<td>72.8</td>
</tr>
<tr>
<td>Coupon</td>
<td>0</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>Short selling fee</td>
<td>0</td>
<td>−3.5</td>
<td>−3.5</td>
</tr>
<tr>
<td>Total</td>
<td>80</td>
<td>89.3</td>
<td>69.3</td>
</tr>
</tbody>
</table>

It is now clear that this trade was done with a bearish view in mind. A less bearish view would have included perhaps a smaller hedge in the stock, maybe by short-selling a number of stocks less than 10.

3.5. Relative Value - Fixed Income Arbitrage. This is a denomination for several different trading styles: Fixed Income-Asset Backed, Fixed Income-Convertible Arbitrage, and Fixed Income-Corporate. Their common denominator is that hedge funds of this type make their returns by realization of a spread between fixed income instrument and other financial instruments. In other words, this strategy trades yield curves. It could be trading credit spreads (high yield/credit arbitrage), mortgage related securities, currency/yield arbitrage, or simply treasury curve relative value trading.

[Example 3.5] shows a simple case of profiting from spreads.

Example 3.5. Suppose now T-bills are trading at 94.2, and Eurodollar futures are trading at 93.1. After some quantitative analysis, an arbitrageur predicts that the Treasury/Eurodollar spread to increase. He buys 10 Eurodollar futures and
10 T-bills. Then the spread indeed increases, with T-bills at 93.95 and Eurodollar futures at 92.7. The arbitrageur gains 40 on the Eurodollar futures and loses 25 on the T-bills, making a net gain of 15.

The best-known hedge fund of this type was Long Term Capital Management (LTCM). It was set up by John Meriwether in the early nineties, together with Robert Merton and Myron Scholes, both Nobel laureates working on option pricing theories. Its trading strategy consisted in trading differences on yield curves across different countries. They looked for pairs of assets whose prices are not in line with each other, and then bought the cheaper one while sold the more expensive one. This distortion in price should diminish in time, with the cheaper one increasing in value and the more expensive one decreasing in value, yielding a convergence in their prices, possibly over a long term time horizon. Then this strategy provides return on both sides once the distortion is eliminated. The strategy turned out to be effective and the fund was successful, having reached several billion dollars in size by 1998.

However, when Russia defaulted on its debt, credit markets dried up causing the yield differences LTCM was trading to widen even more, LTCM began to lose money, margin calls were generated beyond their ability to respond, and eventually went into collapse. [Table 14] explains how the loss occurred.

<table>
<thead>
<tr>
<th>Financial Instrument</th>
<th>Cheaper one</th>
<th>More expensive one</th>
</tr>
</thead>
<tbody>
<tr>
<td>Position</td>
<td>Long</td>
<td>Short</td>
</tr>
<tr>
<td>Source of return</td>
<td>price goes up</td>
<td>price goes down</td>
</tr>
<tr>
<td>Unwind position</td>
<td>Sell</td>
<td>Buy</td>
</tr>
<tr>
<td>Influence on price</td>
<td>decreases</td>
<td>increases</td>
</tr>
</tbody>
</table>

From [Table 14] we see that when the fund gets into trouble and tries to unwind its positions, prices of these pairs of instrument would change in the opposite direction than desired. The value of long position goes down and value of short position goes up, incurring losses on both sides of the trade. This is especially disastrous for major players in the market such as LTCM. Since LTCM had such a big position in these financial instruments, unwinding position is difficult and put much more pressure on the prices. The strategy was set up to benefit from convergence of the instrument prices, but ended up spreading the gap even bigger, to the extent that it is impossible to cover the positions anymore.

From the dramatic fall of LTCM, we summarize some risks that Fixed-Income strategies are exposed to.

- Leverage.

The relative value obtained from a single position in a pair of assets is very small, so big positions and market participation are needed to generate reasonable returns. The funds typically raise a large proportion of the capital by debt, in order to generate substantial returns to the equity holders. In LTCM for example, only $4 of every $100 is from equity and it reached a huge debt to equity ratio of 50:1 at some time points.
leverage can magnify returns when strategy works the way it is designed to be, it is also destructive when things go wrong, with a small loss wiping out the portfolio return.

• Liquidity

As we mentioned above, because of the mere nature of curve trading, notional positions are typically several orders of magnitude the equity capital. For large funds, this can present liquidity problems. More specifically, the funds have counter-parties to each of the financial instruments. The counter-parties execute Mark-to-Market practice on a daily basis, and when the fund suffers from a loss, it would be required to put up more collateral through a margin call. However, when the fund is not performing well, additional capital is harder to obtain, and liquidity dries up quickly. Additionally, as a big player such as LTCM, covering big positions is difficult in the sense that it affects the price significantly and buy/sell trades take longer to be executed.

• Model risk

This relates to the complexity of fixed income pricing, since interest rate markets are difficult to model. Hedge funds operating in this sector are about the most mathematically sophisticated and model risk can affect their operations.

• Credit

One simple and typical strategy in Fixed-Income category is to short Treasury bond futures, while long mortgage-backed or corporate debt securities, taking advantage of higher yields offered by corporations relative to the US Government. The investment is effectively in a “credit spread” and its source of return is the difference in yield curves between the two securities. If the hedge fund performs high yield arbitrage trades, they might be dealing with bonds with of low credit rating or inadequate collateral.

3.6. Event Driven - Distressed Securities. Distressed Securities is an investment strategy that buys a large proportion of debt of a company that is in distress. There are many reasons why a company can be in such a position, such as liquidity shortage, operational shortcomings, legal difficulties, filing for bankruptcy, and reconstruction. Debt is purchased at great discount, and fund managers are typically actively involved in the management and reconstruction of the company. After a period of time when the company is rebuilt, say 1 year or two, the hedge fund sells the debt of the company, when the company distressed is relieved. At that time, the debt would be sold at smaller discount, often at a high premium over the deeply discounted original price, and the hedge fund earns the capital gain from the sale. Source of income for this type of strategy often include the capital gain from sale of the debt, and interest payment coupon payment for owning the debt for a period of time and possibly the sale of the company stock after restructuring.

The next example shows that the rate of return for this strategy can be high, when it is successful:

Suppose a company debt is trading at 50% of par, essentially a bankruptcy liquidation level, which also provides annual coupon of 11%. Now the hedge fund identifies this company as investment target, and buys out the debt of the company,
with a covenant to re-structure the company. The fund manager then is actively engaged in reconstructing the company for one year. A year later, the company obtains a higher credit rating, say BBB, and the hedge fund now sells the debt at 70%.

The ROR for this investment comes from two parts. A capital gain of 40% on the debt, and the coupon payments it received during the year. A total of 51% annual return is obtained.

3.7. Event Driven - Reg-D, PIPEs. Regulation D is a regulation of US Securities and Exchange Commissions that simplifies filing requirements for public companies that sell securities exclusively through a private placement. This financing strategy is also called PIPE: Private Investment in Public Equity. PIPE is an important funding resource for small firms since they have few financing alternatives. In addition, PIPE financing is more efficient than public offerings, since there is no need to approach public investors and develop strategies to sell the shares. Moreover, since the Sarbanes-Oxley act of 2002, imposed in the United States after the Enron accounting scandal, PIPE’s are becoming more common as larger companies also face accounting and filing requirements that make issuing common equity more complex in the US.

(1) How is it done?
As a Hedge Fund strategy, hedge fund managers invest in micro or small cap public firms that issue new restricted shares through private deals. That is, public firms raise money in the private equity market. The shares are sold directly in negotiated transactions, usually at a great discount.

(2) Why is it attractive to investors?
- Discounted Price.
  As mentioned above, issuers in this case offer bigger discount, usually between 5% to 25% of the market value, compensating for the fact that these shares are illiquid before the securities are registered with SEC, and can not be traded publicly during the waiting period. The price of PIPE’s will eventually converge to the price of the publicly traded common shares.
- Liquidity upon registration.
  These shares become public once the firm finishes registration with SEC, and investors can easily sell them at the market price, since the firm is already public.

(3) Hedge a PIPE investment.
Hedge funds that invest in micro-cap PIPEs can hedge their position by short selling the issuer’s stock as the PIPE deal is disclosed. This protects the value of the portfolio no matter the stock price goes up or down after the PIPE deal. Most managers keep this hedged position until the registration is completed and then liquidate the position gradually in the following 90 to 120 days.

(4) Risks involved.
To hedge the position in the short selling position of the firm stock, the hedge fund must be able to borrow securities from shareholders, bank, or prime broker, to ensure the stock can be delivered to the buyer. However, this may not be easy to do since PIPE issuers are small public firms with
small or micro capitalizations. If the shares are not available at time of delivery, "fail to deliver" happens. This is illegal if it constitutes illegal stock manipulation. Short squeezes, therefore, are a major source of risk.

3.8. **Global Macro.** The global macro approach seeks return by taking positions in a wide range of financial instruments, including futures, currencies, indices, commodities, and interest rates. They detect deviation of the perceived value from the actual value by applying macroeconomics principles. Such hedge fund managers forecast in world economies, political developments, and other macroeconomics in their decision process.

Global macro strategy is classified as directional strategy, as opposed to the market-neutral strategies. In a market-neutral strategy, managers seek to neutralize market risks by taking offsetting positions in instruments, based on their theoretical positions. They try to limit risk and exposure to systematic risks, such as changes in macroeconomic variables. Contrary to this, the Global Macro strategy bets on these systematic factors and price movement of global financial instrument.

The most famous global macro fund is probably the Soros Fund Management, run by George Soros. In 1992, the fund took short positions on the GBP against European currencies in 1992, when Bank of England was reluctant to raise its interest rate to levels comparable to other European countries, or float its currency. The fund took such a position because the manager was convinced that the British Pound had to be devalued. Finally, the Bank of England was forced to withdraw from the European Exchange Rate Mechanism and allow its currency to devalue. Soro's fund earned around $1 billion in the trade.

Many of the biggest and most famous hedge funds are Global-Macro strategy based. These funds are generally big, powerful, and aggressive, since they need a large amount of capital to invest around the world and applying all kinds of financial instruments. As in the above example, Global Macro fund can be extremely profitable, when they make aggressive and correct bets about the macroeconomic factors, but can also suffer great losses when the prediction is wrong. Therefore Global Macro strategy is more volatile than others. [Figure12] compares the standard deviation of daily returns of Global Macro strategy to some other strategies including Equity Market Neutral Hedge, Merger Arbitrage, and Relative Value. It is obvious that the volatility of Global Macro funds is significantly higher than others, with a gap of 20% to 40%, which is a great amount of volatility.

3.9. **Managed Futures.**

3.9.1. **Definition.** Managed futures refer to trading in future contracts on both commodities and financial derivatives. Managers following this style are also called Commodity Trading Advisors (CTA), and have a history which pre-dates the rest of the hedge fund universes. The earliest organized futures market dates back to Japan, where an organized futures market for rise was created. Managers that trade exclusively in future contracts have been around for a long time.

3.9.2. **Mechanism.** The futures exchanges provide a transparent, public trading environment, where investors provide initial margin or good faith deposit, and enter future contracts with zero initial values. Investors keep the initial margin account with the exchange, to cover potential future losses. On a daily basis, the Mark-to-Market Margin is in effect, meaning that the current market value of the position and its movement from previous day are calculated and posted on the margin.
account. All traders participate as equals once margin is posted. In case of a loss, it must be topped up, and in case of a profit, the investor receives the gain.

Clearing House is anonymous central guarantor and counter-party of all trades, and they only post their net position on the exchange. This central counter-party eliminates the credit risk and failure to deliver for the actual counter-party of the contracts.

An important warning is that futures prices are rarely a prediction of the price of the commodity in the future. Instead, they are reflections of current market situation, adjusted for interest rates, storage costs of commodities, and seasonal factors.


- Price Certainty
  Futures contracts specifies amount, price, and date of delivery of the asset. Therefore, buy long (short) a future contract, investors are able to lock in a purchase (sell) price, which eliminates the uncertainty of price changes in the future. That is, futures function as an insurance market, provide price certainty to commercial entities.

- Liquid and regulated public market
  Future contracts are standardized and then traded on futures exchange. More specifically, the futures exchanges create uniform products, in which quality, delivery date, and location of product are not variable. The exchanges are organized public markets that preclude special inside information. For example, complete continuous disclosure of prices are offered in 30 different countries. The market is also highly regulated by national governmental agency, such as the Commodity Futures Trading Commission in the US.
3.9.4. Types of Investors. There are two types of investors in the futures market: hedgers and speculators.

Hedgers enter the market and use future contracts to protect themselves against price movements of the underlying asset. For example, a farmer that produces wheat can sell a future contract that specifies the amount, price, grade, and date of delivery of the wheat. This eliminates the risk that wheat price drops dramatically at the time of harvest.

The other type of investors are speculators, who are not interested in physical settlement of the underlying. Rather, they form expectations about futures prices, executive trades on them, and realize short term investment returns. In this case, the traders close out the position and the contract is settled in cash, rather than physical settlement. Speculative traders are very essential to the futures market, since they conduct majority of the trades and that provides trading volume and market liquidity to the futures market.

3.9.5. Benefit of Managed Futures.

- Profit independent of economic conditions
  Managed futures can take advantage of commodity price trends no matter of the market direction. That is, it has profit opportunities in both bull and bear markets, opposed to typical losses in stock or bond portfolio in distressed times. Also, it does not depend on inflation or deflation, nor are they considered foreign property.

  Their best characteristic is their totally different risk profile to other hedge funds. For example, the average manager adopting managed futures strategies had large gains in distressed times such the summer of 1998 and 2002. In our language, managed futures poses negative distress correlation with most other trading styles and provide potential for decreasing portfolio risks. As a result, it has become an increasingly popular addition to global portfolios to improve portfolio quality.

- Free Leverage
  As mentioned above, future contracts are not an asset, have no initial value and only represent a future obligation. Therefore, they offer free leverage, in the sense that it allows investors to take a significant position with a low transaction cost, opposed to trading directly in the commodity. It is also capital efficient since trades are executed on notional terms and cash is used exclusively for margin, while capital can be left in other investment such as T-bills.

3.9.6. Trading Styles. Traders in futures market are Commodity Trading Advisors (CTAs) and there are three main skill sets in this trading strategy category.

- Trend followers
  They are long-term traders that try to extract value from market trends. They are exposed to a high degree of volatility, since “anything can happen in the long run”. Traders of this type develop customized trading algorithms that follow the market trends and subsequently place bets on futures price.

- Short term and counter trend managers
This trading strategy is less volatile compared to trend followers, and make money by taking advantage of short trends. A widely known property of future markets is “mean reversion” (except equities), since price movement create supply/demand changes, which in turn affect the commodities in the opposite way. This property provides source of returns for trading strategies. The traders of this type make frequent, even intra-day trades, and seek profit from trend reversals, such as declining trend from a high point, or increasing trend from a low point.

• Specialty and fundamental managers

This type of traders are driven by economics views, and it is similar to the Global Macro hedging strategy.

4. Portfolio Theory

4.1. Efficient Frontier. The concept of the efficient frontier by Harry Markowitz (see [Markowitz], [Merton]) is one of the fundamental developments of modern portfolio theory.

The Markowitz portfolio theory gives rise to measures for the “expected rate of return” and “expected risk” of portfolios of assets. The theory states that, under a set of assumptions, variance of the rate of return is a useful measure of the portfolio risk. This result of a vast simplification of the portfolio concept, and gives rise to some very essential ideas in portfolio theory.

• Portfolio returns can only be obtained at the cost of taking a risk, which is identified as the standard deviation of returns.

• Portfolio variance depends on the variance of each single asset, but more importantly, the correlation between them. This shows how to effectively reduce the total risk of a portfolio, or diversify the portfolio.

Based on the theory, a portfolio of assets is considered to be efficient if it provides the highest expected return, for a given level of risk. In other words, return cannot be improved without increasing risk, or risk cannot be diminished without hurting the return. In a plane, we can plot expected returns of all the possible investment portfolios against their standard deviation. The set of all these points is called the feasible set, and it can be verified to be convex. Then, the mathematical boundary of this two-dimensional set is denoted as the Efficient Frontier. Intuitively, the efficient frontier represents portfolios that has maximum rate of return for every given level of risk, or minimum risk for every level of return. Efficient portfolios are the portfolios that locate right on the efficient frontier.

The implication of this theory is profound. According to this theory, rational investors should always invest in efficient portfolios. The theory simplifies the investment problem from a two-dimensional decision process down to a one-dimensional one: the efficient frontier is a smooth curve where decisions are to be made, but the investor has the responsibility of making the final decision as to his risk level or return objective to select the appropriate portfolio on the curve. That is, investors target a point on this curve based on his utility function for risk and return.

Using hedge fund data, together with S&P 500 daily returns, we construct the efficient frontier, as in [Figure 13]. In the graph there is also the Sharp Ratio curve of the efficient frontier.
4.2. Sharpe Ratio.

Definition 4.1. Under the assumption that standard deviation captures the entire portfolio risk, and given a risk-free return $c$, Sharp ratio for a portfolio is defined as:

\[ S = \frac{r - c}{\sigma} \]

where $r$ is the mean return of the portfolio, and $\sigma$ is the standard deviation of the portfolio.

Sharpe ratio is the most commonly used tool in portfolio performance evaluation, and we address its importance as well as limitation by explaining the following key aspects.

1. Sharp ratio brings return and risk of a portfolio into one scale-free number, which allows for comparison between different portfolios. Sharpe ratio allows the investor to make an optimal choice on the efficient frontier. Markowitz theory reduced the investment decision from a two-dimensional problem to a one-dimensional one, and Sharpe ratio reduces it further from one-dimensional to a single point.

2. The Sharpe ratio uses the concept of the benchmark; in its simplest formulation, take the benchmark as the risk-free interest rate. $r - c$ represents the mean excess return of the portfolio of the risk-free rate, and therefore the Sharpe ratio measures the mean excess return per unit of risk. Portfolios with higher Sharpe ratio are considered to be better because it provides higher return for the same level of risk taken by investors.
Imagine one is looking for the portfolio that has the best chance of optimizing its performance against a benchmark given by the risk-free interest rate. The following theorem provides a way to do so.

**Theorem 4.2.** Under the assumption that the portfolio return is normally distributed, and given risk-free return \( c \), the portfolio that maximizes the probability of outperforming the risk-free rate is the one with the largest Sharp ratio.

**Proof.** In mathematical form:

\[
Pr\{\Pi > c\} = Pr\left\{\frac{\Pi - \mu}{\sigma} > \frac{c - \mu}{\sigma}\right\}
\]

(4.2)

\[
= 1 - \phi\left(\frac{c - \mu}{\sigma}\right)
\]

(4.3)

\[
= \phi\left(\frac{\mu - c}{\sigma}\right)
\]

(4.4)

where \( \Pi \) is the portfolio return as a random variable, \( \phi \) is the cumulative distribution function of the gaussian, \( \mu \) is the mean return of the portfolio, and \( \sigma \) is the standard deviation of the portfolio.

Based on the important assumption that return is normally distributed, we are able to transform the probability to normal cumulative distribution function \( \phi \). Maximizing this probability is equivalent to maximizing Sharpe ratio because normal cumulative distribution function is a monotonic increasing function.

(4)

Being the most useful tool in modern portfolio theory, Sharp ratio is often misunderstood. The key assumption in Sharp ratio is that standard deviation fully characterizes risk. This assumption is fully satisfied if we have normally distributed returns, and is approximately true with a symmetric return distribution. However, if we notice a highly asymmetric, non-normal return pattern, Sharp ratio can be misleading in portfolio evaluation.

5. The Normality Assumption

In our discussion above, we saw that the Sharpe ratio is easy to understand when returns are normally distributed. Also, Markowitz theory, by virtue of the fact that it only looks at the average return and standard deviation of portfolio returns, is also making an implicit assumption that portfolio returns are normally distributed. This has prompted many authors and practitioners to claim that both theories are useless when viewing hedge fund investments, because of the non-normal nature of their returns, and a collection of alternative theories have been proposed.

A closely related measure of risk is VaR, or *Value at Risk*, introduced by J. P. Morgan in 1994. It measures the largest loss a portfolio can exhibit within a certain confidence level (usually 95% or 99%) and within a given time frame (usually one day). It is easily seen that, if distributions are normal, then VaR, which is simply a loss quantile, will be equal to a multiple of the standard deviation of the portfolio. VaR is a commonly used measure of reserve capital for Banks, but is also often
used by portfolio managers, who are not subject to capital requirements, as an additional measure of risk.

A widely recognized problem with assuming normal distribution is that it does not work well capturing the extreme events, known as “tail events”. For example, the normal distribution says that double-digit losses over a few days should be very rare and only happen much less often than once a century. However, we saw plenty of them in the 20th century, such as the Black Monday market crash in 1987, and global financial crisis in October 2008. This drawback of the normal distribution assumption raises doubts about the use of VaR. 99% VaR would leave out some devastating losses that happen at a frequency far more often than “rare”. For example, if there is a 0.5% chance that the portfolio would lose 100$M, and a 2% chance of losing only 10$M, then the 99% VAR might only be 10$M and the much bigger loss, although less likely, would go unnoticed. These big losses, however, can potentially wipe out the entire portfolio after years of careful accumulation, and are the tail events that risk managers should really be concerned about.

We analyze the normality assumption applied in hedge funds specifically in some more detail.

- **Normal/Boom Market Condition**

  In [Figure 14], we present the histogram of the monthly returns of the HFRI Fixed Income-Convertible Arbitrage Index for the years 2000-2007. During this period, market conditions and returns were stable. Even without big events such as the liquidation of Long Term Capital Management, the graph has a obvious higher density on the left. This indicates a return distribution skewed to the left, which means losses occur more frequently and are deeper than predicted by a normal distribution. This is the widely recognized property of hedge fund returns: a fat tail distribution.

- **Distressed Market Condition**

  In distressed times, it is commonly acknowledged that fund returns can exhibit unusual return patterns, which makes it even less likely to be normally distributed.

  Here we present return distribution graphs for two typical strategies, including Convertible Arbitrage ([Figure15]), and Equity Market Neutral Hedge (Figure16). Along with the density graphs, we also present normal distributions in each graph for comparison. Distressed time is defined as before, as September 2007 onwards.

  From the graphs, we can see that the return distribution for Convertible Arbitrage strategy is very concentrated in the middle, with a much higher spike than the normal distribution. In Equity Market Neutral Strategy, however, the return distribution is very close to normal, and the normality assumption in this case is much more appropriate.

6. Other measures of portfolio risk and return

In Portfolio theory, under the assumption of normal return distribution, we introduced the Sharpe Ratio. When returns are not normally distributed, mean-variance analysis does not describe the return series anymore, and as we have seen in Convertible Arbitrage return in [Figure 14], we need some other ways to evaluate an investment. There are two aspects we need to consider: return & risk. By
combining the two factors, we use risk-adjusted returns. Here we explain some common ones.

6.1. Semi-standard deviation. It characterizes one side of the return distribution, either below or above the mean or a specific target. So it is simply the “downside” or “upside” counterpart of the usual standard deviation. It is used when one is trying to get a feeling as to the asymmetry of the gain/loss distribution.

Gain deviation measures the deviation of portfolio returns from its expected return, taking into account only gains.

Loss deviation is the corresponding deviation when losses only are taken into account in calculating portfolio deviations. Its formula is:

\[
(6.1) \quad \sqrt{\frac{1}{n-1} \sum_{i=1}^{n} (R - r_i)_+^2}
\]

where \( R \) is the benchmark, by taking only the shortfall of return from the target, we incorporate only the losses of our investment.

Problem with semi-standard deviation statistic: although the idea of only calculating downside is appealing, it is not statistically sound. That is, we do not have a sample unbiased estimator of this statistic.

6.2. Sortino ratio. It was introduced by Sortino, van der Meer in 1991 (see, for example, [SvP]), and is the substitute of the Sharpe ratio when one looks only at the loss deviation, instead of looking at the combined standard deviation. Therefore, Sortino ratio punishes downside volatility only, while Sharpe ratio treats
the upside and downside equally. It is agreed that Sortino ratio matches investment risk better, since people are generally not worried about outperformance volatilities. It is defined as:

$$S = \frac{R - T}{DV}$$

Where $R$ is the return of the investment, $T$ is the benchmark, or threshold for return, and $DV$ is the downside volatility.

A common belief is that by not punishing unusual gains, like the Sharpe ratio does indirectly, one maximizes the upside while maintaining the downside. The problem, however, us that it is based on the concept of the semi-standard deviation, and it inherits its statistical deficiencies.

### 6.3. Moments

When we focus on the mean, variance, and correlation of the return distributions, we have implicitly assumed that the return distribution is normal, and can be fully characterized by these statistics. One of the criticisms of using volatilities and correlations as risk measures is the presence of extreme events in portfolio returns, which will go unnoticed in first and second moment calculations. More specifically, according to normal distribution, events such as the ones in 1987, 1995, 1998, etc. should have never occurred. This induces the use of higher moments in attempt to measure large deviations from the mean and capture tail events. Higher central moments are defined as follows.

$$M_k = E[(r_i - \mu)^k]$$
According to this definition, a normal distribution has moments given by 0, 1, 0, 3, etc. The third moment, skewness, is a measure of asymmetry, and the fourth moment kurtosis measures spread. Excess kurtosis, which equals kurtosis minus 3, is more often used, because normal distribution has kurtosis equal to exactly 3.

More precisely, skewness and kurtosis are the standardized third, fourth moments of a distribution, and they are defined as:

\[ S = \frac{E[(r_i - \mu)^3]}{\sigma^3} \]  
\[ K = \frac{E[(r_i - \mu)^4]}{\sigma^4} \]

Their sample estimators are

\[ skew = \frac{1}{s^3(n-1)(n-2)} \sum_{i=1}^{n} (r_i - \bar{r})^3 \]

\[ \kappa = \frac{n(n+1)}{s^4(n-1)(n-2)(n-3)} \sum_{i=1}^{n} (r_i - \bar{r})^4 - \frac{3(n-1)^2}{(n-2)(n-3)} \]

where \( s \) is the commonly used sample standard deviation, defined as \( s = \sqrt{\frac{1}{n-1} \sum_{i=1}^{n} (r_i - \bar{r})^2} \).

To apply the theoretical statistics to real investment data, we present a rolling window estimates of skewness and kurtosis for HFRX daily Index value. We chose
the several typical strategies that we discussed earlier: Equity Hedge, Convertible Arbitrage, Relative Value Arbitrage, Equity Market Neutral Hedge, Distressed Securities, and Global Macro.

Together with these graphs, we can summarize the problems and cautions with using higher moments.

(1) Instability.
Figure 18. 500 day rolling window for HFRX Indices Kurtosis

The moments all show a big jump around November 2008, especially the 3rd and 4th moments are very sensitive to large deviations from the mean. That is, the higher moments of the one or two largest events are so large relative to the other terms in the sum, that they dominate the overall value, leaving other observations irrelevant. This is reflected from the sudden jumps in the data set, even though skewness and kurtosis are
calculated based on a 500-day window, a considerably sized window. This becomes a even bigger problem when we look at specific hedge funds. Many funds have only been operating for a few years, producing a short return history for moments calculation. In this situation, large events cause distortion in the higher moments. In fact, distortion is bigger with higher moment and fewer sample points. For example, [Figure 17] shows the skewness using a rolling historical window, for the monthly returns of a collection of hedge funds. Skewness has a sudden jump from highly negative to positive, when one takes into account just one extreme month, July 2002. This shows the inconsistency of using skewness for insufficient data and it should not be used for serious judgment.

(2) Biased Estimator.

Sample estimators for skewness and kurtosis are biased and can even have opposite sign from the population skewness.

(3) A question we need to ask before applying these statistics to investment decisions is: do moments truly reflect investment returns? We first explain the theoretical support. Positive skewness means a long tail on the right side of the distribution, which translates to frequent small losses and a few extreme gains. This is appealing to investors for two reasons. Firstly, positive skewness means very unlikely extreme downside, and the average positive deviations dominates average negative deviations. Secondly, a positive skewed return distribution has mean higher than median. However, a small example will show that return statistics do not reflect real investment returns precisely, and they can sometimes be deceptive. In [Figure 19], we show the Net Asset Value of an investment, and in Table 15 displays the return summary statistics. It is very attractive when we look at the return statistics: average monthly return of 7%, which compounds to 125% annual return, and a positive skewness of 30.6%, which means there is more outperformance volatilities than underperformance. In reality, this investment loses money, from $1 to $0.2 by the end of the 11th month. The NAV series is shown in Table 16.

<table>
<thead>
<tr>
<th>Table 15.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
</tr>
<tr>
<td>Standard Deviation</td>
</tr>
<tr>
<td>Skewness</td>
</tr>
<tr>
<td>Kurtosis</td>
</tr>
</tbody>
</table>

From this modest example, we should be keep in mind and always be cautious about two things when examining higher moments.

- We should pay close attention to the standard deviation of the investment. When the value is large, such as 0.59 in our example, the moments are most likely to be very misleading. More detailed examination of the investment return, rather than simply the sample statistics, should be carried out to reach a correct investment decision.
- One or a few big events can dominate the moment statistics. In our example, there are two controlling events: a 90% loss from period 7 to
8, and a 100% gain from period 8 to 9. These two events contributed to most of the skewness and kurtosis, leaving other smaller shocks almost irrelevant. Similar to the point above, this means whenever we try to evaluate the quality of investment, focusing only on the moments can be dangerous. Instead, tracking the asset value or cash flow over time can be more reliable.

6.4. Omega. W. Shadwick and C. Keating introduced the concept of Omega a few years ago, as the replacement of the Sharpe ratio when returns are not normally distributed. Their aim was to capture effects of all higher moments fully and use Omega to rank and evaluate manager performance. Their proposition was that, once the fat tail behavior has been captured, one only needs to optimize
investment portfolios to maximize the upside, while controlling the downside. This is sometimes called "risk-adjusted return".

Omega is defined as follows:

\[
\Omega(r) = \frac{\int_{r}^{\infty} 1 - F(x)dx}{\int_{-\infty}^{r} F(x)dx}
\]

\[
= \frac{\int_{r}^{\infty} (x - r)\rho(x)dx}{\int_{-\infty}^{r} (r - x)\rho(x)dx}
\]

\[
= \frac{E[max(R - r, 0)]}{E[max(r - R, 0)]}
\]

(6.10)

\[
= \text{Dembo's reward}
\]

(6.11)

\[
= \text{Dembo's regret}
\]

Where \( r \) is a threshold or benchmark selected by the investors, \( f(x) \) is one-period rate of return on the investment, and \( F(x) \) is the cumulative distribution function of the one-period return. As stated in their paper (see [SK]), all available information from the return distribution, including higher moments, is contained in cumulative distribution function, and hence is encoded in Omega. To provide more intuition about Omega, we show the construction of it using HFRX Equity Hedge Index 2003-2009 data.

We start with the histogram of the index daily return, shown in [Figure 20].

Figure 20. Histogram of HFRX Equity Hedge INDEX Daily Returns 2003-2009
Then we construct the empirical cumulative distribution function of the return series, shown in [Figure 21]. Along with it, we show one benchmark of 50% a year, which corresponds to 0.16% daily return on basis of 252 trading days a year. The two shaded areas are the areas of interest in the calculation of omega. Area 1 is the expected loss, and area 2 is the expected gain, given the benchmark. Omega is simply the ratio of these two areas. In [Figure 22] is the cdf curve for a different hedge fund strategy, Relative Value Arbitrage. Note that the two cdf curves are both very tight, with majority of the return data within a very narrow band around the mean. This is because that hedge fund investment has lower volatility compared to traditional investments.

![Figure 21. Cumulative Distribution Function (cdf) of HFRX Equity Hedge Index Daily Returns 2003-2009](image)

As a statistical observable, Omega presents intriguing possibilities: it measures the proportion of over-performance relative to under-performance, and as such it is conceptually clear. However, we should be cautious when interpreting omega, in the following aspects.

1. Statistical observables for returns do not, in general, mix well with investments performance, just as the previous example of using moments to evaluate performance. In fact, the Omega of our positively skewed investment earlier is larger than 1 (for most benchmarks), despite the fact that it loses money.

2. As a ratio, Omega is the ratio of two first order partial moments of the return distribution, and carry little information about volatility, a second moment. Therefore, unlike Sharp ratio, omega does not penalize for high volatility of an investment. In fact, in the paper by Gilli et al. ([GSDC]), it is shown that if we construct portfolios by using Omega function, they exhibit higher volatility.
(3) Omega function is a curve, not a number, and it is most useful when considered as a function of the benchmark return \( r \). Indeed, it is not a point estimator, but rather it considers the entire shape of the distribution with respect to the benchmark. When used in risk management, the benchmark should be chosen to be low, or even a negative number with large absolute value. In this case, a larger Omega ratio indicates a safer investment. However, by choosing investments with large omega in this case, it provides no information about the probability of over-performance and obtaining attractive return. On the contrary, if we use a medium-sized benchmark, portfolio return objectives are considered, but negative tail effects are ignored. Therefore we conclude that when using Omega ratio, we need to judge the ratio against the benchmark chosen, and balance between low risk and high return. For hedge fund portfolio optimization, it is not clear which benchmark to use in the optimization process, and some kind of Omega benchmark utility theory is needed for decision making.

7. Correlation risk

7.1. Correlation between hedge funds. Hedge funds employ a wide range of financial instruments to reduce volatility while enhancing expected return, and have also displayed low beta relationship with the traditional stock market. [Figure 23] tracks the correlation between HFRI Equity Market Neutral Index and SP500, based on historical monthly returns. Positive correlation does not exceed 0.3, and in half of the years, correlation is negative. Another typical example is for Fixed Income Strategy, shown in [Figure 24], which also exhibit low or negative correlation with the equity market. These figures show that hedge funds constitute excellent
diversification, which is the main property that hedge funds bring to investment portfolios.

**Figure 23.** Correlation between Equity Market Neutral Index and S&P 1990-2007

![Correlation between Equity Market Neutral Index and S&P 1990-2007](image)

**Figure 24.** Correlation between Fixed Income Index and S&P 1993-2007

![Correlation between Fixed Income Index and S&P 1993-2007](image)

Exactly because diversification is a key attribute of hedge funds, situations in which this diversification disappears are considered to be grave.

7.2. **Correlation breakdown.**

7.2.1. *Standard assumption.* The standard assumption when studying asset returns is that returns have the multivariate Gaussian distribution. This assumption can be reasonable and permits the construction of tractable and useful models.

7.2.2. *Loss of normality.* When we look at returns of many funds simultaneously, normality can be lost for two reasons.

- Marginal distribution may not be normal.

As we discussed in section “normality assumption”, individual assets frequently have richer, non-Gaussian features, such as skewed (asymmetric around the mean) and leptokurtotic (“fat-tailed”) nature. Specifically,
when we presented the performance histogram of the HFRI Fixed Income-Convertible Arbitrage Index, we noted that there are too many large losses that cannot be explained by a normal distribution.

- Dependence structure complication.
  Dependence structure of multiple assets may not be determined uniquely by the correlation, and we call this “correlation breakdown phenomenon”.
We will address this problem in detail, presenting a methodology to analyze correlation risk.

To understand types of events that cannot be reconciled with a normal distribution, imagine an experiment with one hundred fair coins, each with a 50% probability for head or tail. Now flip all of them simultaneously several times. Imagine in all the observations, approximately 10% of the times, all of the 100 coins give tails at the same time. In this case, there is nothing abnormal about each coin’s equal chance to generate head and tail, what is non-normal is their dependence structure. That is: to generate such unlikely events, the tail outcomes of the coins have to be positively correlated with each other in certain situations: when one shows tail, others are also more likely to show tail.

Something similar happens with hedge fund events. In addition to their own individual performance and possibly a fat-tailed behavior, there are months in which too many funds exhibit the same behavior. In May 2005, the vast majority of hedge funds lost money, and something similar happened in April 2004. We will of course also remember the distress in hedge fund market in 1998 and 2008. In other words, mutual behavior between the hedge funds is something that according to a normality assumption should only occur once in a hundred years.

7.2.3. Solution to correlation breakdown. We will investigate these facts and problems from a regime switching perspective. Regime switching models describe process in which parameters of the process change discontinuously according to a realized stochastic path. In our case, we will distinguish between two states of the market: normal and distressed. Between these two states, the parameters of our models have a sudden jump. In following context, we will refer to the normal distribution as Gaussian, to avoid confusion between the two market states.

Mixtures of Gaussians are standard in the study of fat-tailed distributions in the univariate case. Multivariate Gaussians are relevant for our study because they offer a framework to understand how uncorrelated returns can become correlated under distressed market conditions.

An \( n \) -variate probability distribution is a mixture of two Gaussian distributions when the value of its density function \( f(x) \) at an \( n \)-dimensional vector \( X \) can be expressed as a linear combination of two Gaussian densities as follows:

\[
pe^{\frac{1}{2}(X-M_1)'A_1(X-M_1)} + (1-p)e^{\frac{1}{2}(X-M_2)'A_2(X-M_2)}
\]

Here, \( M \) denotes each of the two means in each of the two states, and \( A \) denotes each of the correlation matrices. This models a process where each event in a time series is normally distributed, but it follows one or the other distribution with probability \( p \) or \( (1-p) \) respectively. We refer to \( p \) as the mixing coefficient. Or alternatively, events are normally distributed when we are in either of the states, but there are possible switching from one state to the other. The two normal distributions in different states have different parameters.
This distributional model, with its associated portfolio optimization framework, was developed in [BSS].

7.2.4. Estimate the parameter. To work with such a regime-switching mixed distribution, it is necessary to calibrate using market data, in order to find the parameters (means, variance-covariance matrices, mixing coefficients) that uniquely define it.

A better procedure is to use maximum likelihood estimators, or Bayesian methods through Markov-chain simulation. However, the drawback of this approach is that the resulting distribution offers no insight as to what constitutes distress.

For the purpose of this article, we will offer a simple alternative that explains market behavior, while remaining compatible with maximum likelihood estimators. In particular, we will use this to model normal situations, instead of market distress. This is done as follows.

Define distress in terms of the deviation of a particular hedge fund from its mean in either direction. More precisely, we say that a hedge fund is in distress in a given month when its performance deviates from its mean $\mu$ by more than a given number $n$ times the standard deviation: i.e. $|r - \mu| > n\sigma$. $n$ is frequently chosen to be 2, because it is simple and also meaningful as events more than two standard deviations away from the mean should be practically impossible according to a Gaussian rule. Furthermore, we will call a certain month distressed when sufficiently many funds are in distress. One would find that the hedge fund universe is in distress about 25% of the time.

Note also that our definition of distress allows us to account for exceptionally bad months as well as exceptionally good months (such as May 2003). The opportunistic investment reader may be discouraged by this, as it appears to be something that may result in confusing good and bad performance. Our ultimate objective, however, is very different: we just want to be able to identify risk convexity, which only manifests itself in large event situations (positive or negative), hence both type of events, good and bad, are going to be very useful in determining risk commonality throughout the hedge fund universe.

To understand the value of this approach, consider the following exercise: take a large collection of hedge funds (say 30) and measure their correlation during the normal times and during distressed times separately; what we get is that the correlation during normal times is given by Figure 25.

Here we are adopting a pictorial approach to correlation matrix display: to each number in the square array, we attach a high pile when it is close to +1, and a low pile when it is close to -1. This allows us to visualize a matrix easily. When we look at the distressed correlation matrix, we get instead Figure 26.

What we see here is very interesting: many of the correlations are driven higher, close to $+100\%$ during times of market distress; but a handful of them actually become lower, identifying in this manner hedge fund which help with diversification in the worst of times. We could take this pictorial approach a little further, turning each element in the matrix into a pixel with some coloring rules; this would allow us to summarize both matrices by a combined 2-dimensional aerial view, which we can use to quickly display the correlation properties of even larger collections of hedge funds.

The conclusions are the same in both situations: while the correlation numbers are fairly low during normal market conditions, they spike to +1 during market
distress, except that a handful of managers actually gives us even lower correlation during market distress; for the cautious investor, these managers bring a type of investment insurance with them that, if their individual returns and risks are appropriate, may make them very desirable for a balanced portfolio. What all of this is pointing to is a need for a Markowitz theory that takes into account regime switching. We refer the interested reader to the article by Buckley, Saunders and Seco ([BSS]) for a mathematical treatment of these issues.

7.3. Correlation breakdown effect for CFO’s. Recall our previous discussion on Collateralized Fund Obligations using the case of Diversified Strategies CFO SA, launched in 2002. Investors provided equity worth $66.3 million, which supported an investment of $250 million in hedge funds. The additional funds ($183.70 million) were raised through three bond tranche issues, as follows:

- AAA tranche ($125 million)
- A tranche ($32.5 million)
- BBB tranche ($26.2 million).

We are not going to go in great detail into the details of the transaction; we will simply mention that the tranche structure is similar to a CDO; the bond investors provide the capital, and upon maturity, get their principal and interest. In case, the CFO structure fails to have enough assets to pay back its debts, the CFO will enter into default. In that scenario, the AAA-tranche investors are first in line to get their money back (principal plus interest). Next in line will be the A tranche,
and the BBB tranche will be the last in line. In a default situation, the equity investors would have lost all their assets. Because of the difference in default risk, each of the bond investors receive different interest payments, highest for the BBB tranche, lowest for the AAA investors. The interest payment for what their risk is worth a credit spread— is a very interesting risk pricing problem, which can easily be solved in the case of gaussian distributions for the underlying hedge fund assets. It is easier than the traditional CDO pricing problem, since here we only need to look at the performance of the entire fund performance, and we do not need to enter into individual default numbers. In fact, with the assumption that the fund returns are normally distributed, it is very easy to determine the credit spread. The probability of default will be given by the quantile of a normally distributed Itô process, which has a simple risk-neutral analogue, and we just price that using expectation under the risk neutral measure. In the case of the Diversified Strategies CFO, the respective interest rates were as follows:

- AAA tranche: LIBOR + 0.60%.
- A tranche: LIBOR + 1.60%.
- BBB tranche: LIBOR + 2.80%.

However, if we now assume the underlying hedge funds have a mixture of multivariate gaussian distributions, in accordance with the discussion in our previous section and presented in [BSS], the pricing problem becomes an entirely different matter.

We are going to obviate the details, which the interested reader can find in [AEHS], and simply summarize a rather striking result of how the default probabilities depend on the distress parameter \( p \) in 7.1. Figure 27 displays the default

**Figure 26. Distressed correlations**
probabilities of the different tranches of the CFO, where the parameter $p$ is allowed to vary from 0 to 1. We see how the probability of default jumps to dangerous levels as $p$ approaches 1. With the perspective we have now developed from the events of 2008, this picture is perhaps not so surprising, but it shows how distressed considerations can help foresee risks not apparent in gaussian models.

**Figure 27. Default Probabilities with different correlations**

8. Manager default risk

8.1. Examples of business risk. Business risk in the hedge fund industry is synonymous with operational risk, and it refers to everything that isn’t market or credit risk. The best-known business risk event is the blow-up of Long Term Capital Management in 1998, since this event not only involved the investors in the fund, but perhaps was even more dramatic for the financial system as a whole, given the size and the number of players that were involved in it. The events started when Russia defaulted on its debt, and investors’ confidence was shaken due to President Clinton’s impeachment possibility. The result was a massive credit and liquidity crunch in the market. LTCM was forced to unwind its position to cover losses and deliver money to cover margin calls. However, since LTCM had such a big market share, unwinding was very difficult and market prices were strongly influenced. This led to the freezing of LTCM’s assets, since its positions were largely uncollateralized. It was a nightmare scenario for the design of the BIS-II resolutions on credit risk management.

While not of the same size, events such as this one occur every year: Silvercreek (brought down by Enron’s demise), Lancer (fraud), Beacon Hill (fraud on the part of the managers to attempt to survive massive refinancing orders in the mortgage market), and the Appalachian illegal dealings in mutual fund timing are a few examples of situations where investors lost money for reasons other than normal market fluctuations. The most recent example is the Ponzi scheme by Madoff, the biggest fraud committed by a single person in finance history. In terms of hedge fund, “blows-up” typically refers to a situation where losses occur and law
enforcement acts; the result is always forensic investigations, fund liquidation, and press coverage.

8.2. Model the risk. Modeling this risk in an investment portfolio is critical. All fund-of-fund managers contain a “due diligence” team that ensures all the underlying hedge fund investments are sound. Equipped with a good nose and a magnifying glass, they visit the hedge funds, analyze their positions, follow trade tickets, ascertain independence of the administrators and true asset valuation processes, request audits, and sometimes hire private investigators to look for dark spots in the history of the firms and its key employees. This exercise will be repeated often, annually or perhaps quarterly depending on the management’s requirement. If the result of all the investigation is good, the investment manager approves the hedge fund, and an investment may take place.

In order to get a sense as to how to go beyond the due diligence process, and indeed to build the due diligence process into the portfolio construction process of a hedge fund portfolio, we illustrate an example.

8.3. An example. Assume the entire hedge fund universe consists of 100 managers, and every year, 1 out of the 100 will have a blow-up event (fraud or any other), which will lead to a total loss of the investment in that one manager (no recovery upon loss). Assume also we have a $100 M portfolio invested in hedge funds.

Consider two alternatives:

• A portfolio that invests in all the money in one manager alone.
• A portfolio that invests equal amounts of money in every single manager ($1M per manager)

In deciding what to do, a risk averse investor will not go with the first choice. It is better to diversify and invest in the second choice, knowing that the loss is guaranteed to be $1M (a diversification premium, or insurance cost), and knowing that the portfolio will produce positive returns as long as the average return for the other managers is in excess of 1% per year.

The portfolio that invests equal amounts of money in every single manager ($1M per manager) has the following characteristics:

• The probability of having one credit event inside the portfolio is 100%; it is certain that one of the funds will have a credit event, by assumption.
• The expected loss due to a credit event is equal to the loss that we know we will get for sure, which is $1M.
• The standard deviation of losses is 0; there is no uncertainty as to how much the portfolio will lose: it will lose $1M.

The portfolio that invests all the money in one manager alone has the following characteristics:

• The probability of having a credit event inside the portfolio is 1/100. It is smaller than the previous portfolio, and in fact, holding only one hedge fund, instead of any other number, provides the smallest probability of a blow-up event.
• The expected loss due to credit events is equal to $1M. This expectation is the same as the first, and is in fact the same no matter how many hedge funds we hold in the portfolio.
The standard deviation of credit losses is about $9M.

The conclusion of this example is clear: the standard deviation of losses is the key indicator of the business risk. The probability of blow-up, or the expected loss carries no information, as we saw in the example that these characteristics can be identical for two different strategies. We will propose a methodology to incorporate a manager rating system into a portfolio construction process.

8.4. Malfunction frequency and severity. CreditMetrics is one of the two main methodologies for credit risk management. It is based on a rating system for each counter-party, from which a probability of default is derived. For hedge fund portfolio management, an analog can be established, as follows: First, we introduce a rating system for hedge funds. This rating system will aim at quantifying the due diligence process we described earlier. Let's consider a simple version consisting of only two dimensions: one for blow-up frequency, another for severity, as shown in Table 17:

According to this, ratings will be pairs of letters, such as AA for the best, and CC for the worst. Next, we associate a frequency distribution to the vertical silos and a severity distribution to the horizontal ones, very much like CreditMetrics. Unlike CreditMetrics, we won't have history to calibrate these numbers, and they will be the responsibility of the due-diligence team to determine. Let's just say that, for a given fund, the blow-up frequency is given by a number $\lambda$ and the severity is given by a random variable $\Sigma$, of which we can generate samples denoted by $\Sigma_i$. If the hedge fund has a return history given by $S_i$, we calibrate it to our favorite distribution and generate a sufficiently large future distribution of returns, which we will also denote by $S_i$ (except they could span one hundred years into the future, if we so wanted). After this, we select $\lambda_i\%$ of those return numbers, and we simply replace them by $\Sigma_i$. The result is a modified future return distribution that takes into account possible manager malfunction. One then optimizes the portfolio construction using this generated data instead of the historical one. The resulting portfolio will be one that takes into account market and blow-up risks simultaneously, and is an interesting way of extending Markowitz theory to this context.

References

HEDGE FUNDS


DEPARTMENT OF MATHEMATICS, UNIVERSITY OF TORONTO, TORONTO ON M5S2E4, CANADA

Current address: Department of Mathematics, University of Toronto, Toronto ON M5S2E4, Canada

E-mail address: seco@math.toronto.edu
This page intentionally left blank
Abstract. Credit derivatives often show an appealing risk-return profile for investors. They allow for an enhancement of portfolio returns, but also offer high potential for risk diversification, due to the correlation structure of their returns to those of traditional asset classes. Beside corporate bonds, we therefore review popular credit derivatives and present a mathematical definition of their payoff structures. To actually price credit derivatives, we discuss the most important models and theoretical developments of the past years. Structural-default models assume that the dynamics of the firm’s asset value over time can be described as a stochastic process and that the defaultable security can be regarded as a contingent claim on this value. Intensity-based models do not consider the relation between default and asset value in an explicit way. They rather specify the default process exogenously and model default as the stopping time of some given hazard-rate process. While reduced-form models have attractive properties, their main drawback is the missing link between economic fundamentals and corporate defaults. Hybrid models try to overcome this shortfall and combine the advantages of structural and intensity-based models. Correlated changes in revenues and costs of different companies usually influence the default probability of these firms. We will thus present different approaches for the modeling of joint defaults and give some algorithms to show how these approaches can be implemented for the pricing of portfolio-credit derivatives.

1. Introduction

Until the credit crisis, the credit market has experienced an enormous growth over the last years. After a standardization of credit-derivative contracts and the introduction of credit default swap (CDS) indices, a revolution in terms of outstanding notional has taken place, which is only one reason why credit instruments are very attractive to investors. Credit instruments, such as corporate bonds or credit derivatives, often have an appealing risk-return profile allowing to enhance portfolio returns. Furthermore, due to the correlation structure of their returns to those of traditional asset classes, such as stocks and government bonds, they offer high potential for diversification. A great variety of credit-pricing models have been developed over the last thirty years to model and price credit derivatives. These models can be divided into four main categories. The first and second generation of

Key words and phrases. Credit derivatives, structural-default model, reduced-form model, hybrid model, CDS, CDO.
structural-default models, the reduced-form models, and the hybrid models. The first generation of structural-default models goes back to [59], who basically uses the Black-Scholes framework to model default risk. He assumes that a firm defaults if the market value of the firm falls below the value of its liabilities at some fixed date. Following this basic intuition, it is easy to derive an explicit formula for the price of a defaultable zero-coupon bond. Later models try to relax some of the unrealistic model assumptions of Merton’s framework, e.g. [5], [25], and [50]. The second generation of structural-default models assumes that default may occur any time between issuance and maturity of the debt and relaxes several of the original assumptions. Default is triggered as soon as the value of the firm’s assets reaches a lower threshold level. Examples are the models of [5], [43], [53], [11], [67], [62], and [84]. Despite the immense efforts devoted to generalize Merton’s methodology, most of these structural models have only limited success in explaining the behavior of prices of debt instruments and credit spreads. It is well documented that structural models relying on continuous processes generate unrealistic low yield spreads for investment-grade debt, especially for debt of short maturities (see, e.g., [32] and [24]). In addition, [24] show that the predictive power of continuous structural models is limited. Moreover, while most structural models predict a humped shape for the term structure of credit spreads, [52] find evidence for increasing spread structures, [68] for negative slope, and [30] for positive slope in case of speculative grade debt. These problems have led to attempts to use models that make more direct assumptions about the default process. These so-called reduced-form models define default as a stopping time of some given hazard-rate process, i.e. the default process is specified exogenously. Hence, reduced-form models can be applied to situations where the underlying asset value is not observable and the behavior of credit spreads for short maturities can be captured more realistically. Some important papers in this area of research are [50], [38], [47], [48], [20], [21], and [75]. Empirical evidence concerning reduced-form models is rather limited. [22] finds that these models have difficulty in explaining the observed term structure of credit spreads across firms of different credit-risk qualities. [16] consider specific Vasicek- and Cox, Ingersoll, Ross (CIR)-type reduced-form models and show that both models fail to account for all observed shapes of the credit-spread structure.

There is a heated debate which class of models - structural or reduced-form - is most adequate (see, e.g., [35]). [36] compare structural and reduced-form models from an information-based perspective. They claim that the models are basically the same - the only distinction between the two model types is not whether the default time is predictable or inaccessible, but whether the information is observed by the market, as it is assumed in structural models, or not, as in reduced-form models.

Hybrid models try to combine the ideas of structural and reduced-form models to eliminate their drawbacks. Basically, these models can be seen as a variant of reduced-form modeling with state variables. The conditional probability of default is directly related to specific macro- and/or microeconomic factors. Microeconomic factors can contain firm-specific structural information. Important examples are [13], [56], [15], [57], [72], [39], [2], and [74]. Some of these authors have parameterized the instantaneous credit spread as a function, usually affine, of candidate economic and firm-specific state variables and then directly estimated the effects of these variables.
A young field of research is the modeling of dependent default events. Modeling credit risk from a portfolio perspective is required if derivatives on credit portfolios are priced and if risk measures of credit portfolios have to be computed. Derivatives on credit portfolios have seen an enormous growth over the last decade since new securitization techniques have been developed. Also, new regulatory rules make it mandatory for financial institutions to extend their risk measures from a single-contract perspective to a portfolio view. Most of today’s portfolio models are based on existing single-firm models and introduce dependence by assuming stochastically dependent variables for the individual firms.

This article is organized as follows. In Section 2 we present the most important credit instruments, these are corporate bonds, CDS, collateralized debt obligations (CDOs), portfolio CDS, and nth to default swaps. Section 3 is dedicated to the most important representatives of the first and second generation of structural-default models, while reduced-form models are explained in Section 4. In Section 5 we discuss some hybrid models and in Section 6 we present different approaches for the modeling of joint defaults. Finally, Section 7 concludes.

2. Products and markets

2.1. Bonds and related instruments. Bonds are debt securities which are traded on fixed income markets. The bond’s issuer owes the holder of the bond a debt, which has to be repaid based on a pre-specified payment schedule. In the case of a typical coupon bond, premium payments (called coupons) are paid periodically and the principal (or face value) of the bond is paid back at maturity. Exceptions are zero-coupon bonds, i.e. bonds without coupons but a single payment at maturity, and consol bonds, i.e. coupon bonds that do not mature. Another variant are floating-rate notes, these are coupon bonds with time-varying coupons at a level which is linked to some benchmark return. Bonds are often classified according to their issuer. The vast majority of bonds is issued by one of the following parties: National governments (government bonds), provincial, state, or local authorities (municipal bonds), and private corporations (corporate bonds). While government bonds, such as American Treasure notes, are usually considered to be default-free investments, corporate bonds and some municipal bonds are subject to the possibility of credit default of the issuer. Assessing the probability of credit default becomes crucial when the fair price of a bond has to be determined, as the holder of a defaultable bond demands for a higher interest for bearing default risk. The additional interest, compared to default-free investments, is called credit spread and is often interpreted as the markets view on the creditworthiness of the respective issuer.

If a company defaults, its management loses control and the remaining assets are liquidated. The revenues of the liquidation process are then distributed pro rata, if no priority rule was specified, according to the invested principal of each bondholder. The recovered fraction is referred to as recovery rate and is abbreviated as $R$. Let us briefly remark that different modeling approaches are feasible to describe $R$. These recovery schemes differ in complexity (deterministic versus stochastic $R$) and in the timing when the recovered fraction is paid to the bondholders. The scheme fractional recovery of face value assumes a payment in the amount of $R$ times the bond’s face value immediately at the time of default. The scheme fractional recovery of treasury value defers the same payoff to the maturity
of the bond. Fractional recovery of market value implies that bondholders receive some fraction of the bond’s market price (prior to default) immediately at the time of default. This assumption is often convenient in intensity-based models. Finally, the scheme fractional recovery of firm value assumes the recovery rate to be a function of the firm’s assets and liabilities at the time of default.

The growth of the global bond market over the last years is impressive, compare Figure 1. According to the research report\(^1\) (Nov. 2007) of the SIFMA Securities Industry and Financial Markets Association, the outstanding nominal of governmental bonds and private corporations (separated in financial and corporate) exceeded 67,000 billion US$ in 2006 and is expected to rise further in 2007.

![Figure 1. Growth of the global bond market.](image)

For a mathematical analysis of bond prices it is important to observe that each coupon bond can easily be replicated by a suitable linear combination of zero-coupon bonds with weights and maturities according to the payments of the respective coupon bond. Therefore, we concentrate without loss of generality on the pricing of zero-coupon bonds. Let us remark that the converse, i.e. separating coupons from final payment, is called stripping of a bond and is also common in practice. To simplify the notation we standardize the final payment to one unit of the respective currency. Uncertainty in all models is described using a filtered probability space \((\Omega, \mathcal{F}, \mathbb{P})\), where \(\mathcal{F} = \{\mathcal{F}_t\}_{t \geq 0}\) is the investors filtration and \(\mathbb{P}\) is a given pricing measure. For the sake of simplicity, we assume that \(\mathcal{F}_0\) is trivial and that there exists an \(\mathcal{F}_t\)-measurable short-term interest-rate process \(r = \{r_t\}_{t \geq 0}\). The time of default of the considered firm will be denoted by the random time \(\tau\), for which we will present different mathematical models in the following sections. Based on these assumptions, the price at time \(t\) of a default free zero-coupon bond with maturity \(T\) is given by (for more details on the pricing under a given pricing

\(^1\)http://www.sifma.org/research/pdf/RRVol2-10.pdf
measure, see e.g., [83])

\[
P(t, T) = \mathbb{E}_t \left[ e^{-\int_t^T r_s ds} \right].
\]

Often, the required computations are significantly simplified if a deterministic and flat interest rate \( r > 0 \) is assumed, \( P(t, T) \) then equals \( \exp(-r \cdot (T-t)) \). Within this setup, the fair price of a defaultable zero-coupon bond is given as the expectation of its discounted payoff with respect to the pricing measure \( \mathbb{P} \). If the company survives until maturity, the payoff of the zero-coupon bond is one unit of currency. Otherwise, we assume the investor to receive the fraction \( R \) at the time of default, i.e., we assume a fractional recovery of face value. This leads to the following lemma.

**Lemma 2.1 (Price of a defaultable zero-coupon bond).** Let \( P^d(t, T) \) denote the fair price of a defaultable zero-coupon bond at time \( t \) with maturity \( T \). Moreover, let \( r > 0 \) be flat and known. If we assume a fractional recovery of face value, we have as long as \( \tau > t \):

\[
P^d(t, T) = e^{-r \cdot (T-t) \cdot \mathbb{P} (\tau > T|\mathcal{F}_t)} + \mathbb{E}_t \left[ e^{-r \cdot (\tau-t) \cdot R \cdot 1_{\{t<\tau\leq T\}} \left| \mathcal{F}_t \right.} \right].
\]

The credit spread corresponding to \( P^d(0, T) \) is denoted by \( S_T = S(0, T) \). It is defined as the real number that solves the equation

\[
P^d(0, T) = e^{-(r+S_T) \cdot T} = e^{-S_T \cdot T \cdot P(0, T)}.
\]

2.2. Credit default swaps (CDS). By outstanding nominal and activity, today’s most important credit derivatives (on individual firms) are CDS contracts. Heuristically, CDS can be interpreted as insurance policies against the default risk of some reference entity, which is typically some corporate bond. More precisely, CDS are contracts between a protection buyer and a protection seller, whereby the protection buyer makes periodic premium payments over a predetermined time horizon. This part of the contract is usually called the premium leg. In return, the protection seller commits to make a payment in the event of a credit default of the reference entity. This agreement implies that the term structure of default probabilities of the reference entity is the crucial factor in pricing CDS. One reason for becoming so popular is that CDS allow to directly buy or sell default risk. For instance, CDS are often used to build short positions in credit risk. Moreover, CDS are interesting building blocks for different portfolio strategies and complex credit derivatives. To a large extent, CDS are traded over-the-counter and allow for different contractual variants with respect to the definition of default and the payment at default. While a comprehensive introduction to such issues is presented in [7], we will focus on the mathematical modeling here.

Prices of CDS contracts are usually quoted in terms of an annualized premium. This premium is measured in basis points (bp), where one bp is 0.01%. The fair premium is chosen such that the expected discounted payoffs of both contractual parties agree, i.e., the present values of premium and default leg are equal. This choice allows both parties to enter the contract at zero cost, i.e., without an upfront payment. To begin with, let us assume a known recovery rate \( R \), a flat interest rate \( r > 0 \), and a unit nominal for the contract. Moreover, the default payment
is immediately settled at $\tau$ and the contract matures in $T$. Given the premium-
payment schedule $0 < t_1 < \ldots < t_n = T$, the present values $EDPL$ of the premium leg and $EDDL$ of the default leg at time $t = 0 < \tau$ are given by

$$EDPL = \mathbb{E} \left[ \sum_{k=1}^{n} e^{-rt_k} \cdot s^{CDS} \cdot \Delta t_k \cdot 1_{\{\tau > t_k\}} \right]$$

$$EDDL = \mathbb{E} \left[ (1 - R) \cdot e^{-rt} \right] = (1 - R) \int_{0}^{T} e^{-rt} d\mathbb{P}(\tau \leq t),$$

where $t_0 = 0$, $\Delta t_k = t_k - t_{k-1}$, and $s^{CDS} = s^{CDS}_T$ is the annualized spread of the CDS with maturity $T$. A common simplification is to assume the insurance buyer to continuously pay the spread $s^{CDS}$. In this case, the sum in the premium leg is replaced by the integral $\int_{0}^{T} e^{-rt} \cdot s^{CDS} \cdot 1_{\{\tau > t_k\}} dt$, which is often more convenient to work with. Based on this simplification we can solve for $s^{CDS}$ and obtain for the fair CDS spread of a contract maturing in $T$

$$(2.4) \quad s^{CDS} = \frac{(1 - R) \cdot \int_{0}^{T} e^{-rt} d\mathbb{P}(\tau \leq t)}{\int_{0}^{T} e^{-rt} d\mathbb{P}(\tau > t) dt}.$$

### 2.3. Collateralized debt obligations (CDOs)

CDOs first appeared around 1990 and have grown impressively since then. CDOs are credit derivatives with payment streams that depend on a portfolio of defaultable assets. Hence, CDOs belong to the class of asset-backed securities. The original idea behind CDOs is to pool credit-risky assets, such as corporate bonds, commercial loans, and mortgages, and to resell the resulting portfolio in several tranches with different seniority and return characteristics. A popular variant are synthetic CDOs, where the reference portfolio contains CDS contracts instead of traditional credit instruments. The advantage of this construction is that the issuer of the CDO is not required to own the securitized assets. A capacious introduction to legal issues, taxation, and accounting questions related to CDOs has been published by *JP Morgan*, compare [54].

CDOs are often classified, according to the purpose of the issuing party, in balance sheet and arbitrage CDOs. Balance sheet CDOs are issued by commercial banks or insurance companies to transfer some risk from their balance sheet. Arbitrage CDOs simply aim to profit from reselling the securitized assets to investors who search for a specific risk profile to be found in one of the CDO’s tranches.

From a financial point of view, CDOs are of special interest as their price is determined to a large extent by the correlation among the pooled companies. Pricing the different tranches of a CDO therefore requires a realistic model for the dependence among the pooled assets, as the credit risk of each tranche is affected by idiosyncratic default probabilities (including the default severity) and the default correlation. While markets have developed a deep understanding for the description of individual default probabilities and single-name derivatives, there is no market consent on a model for the dependence structure and the pricing of CDOs until now.

The most subordinate tranche of a CDO is called equity tranche. This tranche sustains all credit losses as long as the overall portfolio loss does not exceed its
nominal. Then, the second tranche bears the losses of the portfolio, etc. This construction of seniority is sometimes illustrated as a cascading waterfall. Investors of a tranche receive periodic premium payments, typically quarterly, until the maturity of the CDO. The amount of each premium payment depends on the remaining nominal of the respective tranche at the premium payment date. In return, they have to compensate for all losses affecting the tranche they are invested in. In what follows, we describe the payment streams of a CDO in mathematical terms. For this, let $I$ denote the number of pooled assets, indexed by $i$, and $J$ the number of tranches, indexed by $j$. The time of default of asset $i$ is denoted by $\tau^i$, its recovery rate by $R^i$, and the overall portfolio loss (as a percentage of the overall nominal) at time $t$ by $L_t$. If identical recovery rates $R = R^i$ and identical portfolio weights of $1/I$ are assumed, then

$$L_t = \frac{(1 - R)}{I} \cdot \sum_{i=1}^{I} \mathbf{1}_{\{\tau^i \leq t\}}.$$  

Tranche $j \in \{1, \ldots, J\}$ is specified by its upper attachment point $u^j$ and its lower attachment point $l^j$, $l^j \leq u^j$. For instance, the first tranche usually covers [0%, 3%] of the portfolio losses. Using these attachment points we can express the loss $L^j_t$ affecting tranche $j$ up to time $t$ as a function of the portfolio loss $L_t$ via

$$L^j_t = \min\left(\max\left(0, L_t - l^j\right), u^j - l^j\right).$$

Today’s market standard is defined by the European iTranx portfolio and its American equivalent DJ CDX. Table 2.3 lists their respective segmentation.

<table>
<thead>
<tr>
<th>$I = 125$ companies</th>
<th>iTranx</th>
<th>DJ CDX</th>
</tr>
</thead>
<tbody>
<tr>
<td>$j$ Tranche</td>
<td>$l^j$</td>
<td>$u^j$</td>
</tr>
<tr>
<td>1 Equity</td>
<td>0%</td>
<td>3%</td>
</tr>
<tr>
<td>2 Junior mezzanine</td>
<td>3%</td>
<td>6%</td>
</tr>
<tr>
<td>3 Senior mezzanine</td>
<td>6%</td>
<td>9%</td>
</tr>
<tr>
<td>4 Senior</td>
<td>9%</td>
<td>12%</td>
</tr>
<tr>
<td>5 Super senior</td>
<td>12%</td>
<td>22%</td>
</tr>
<tr>
<td>6</td>
<td>22%</td>
<td>100%</td>
</tr>
</tbody>
</table>

Table 1. The iTranx and DJ CDX segmentation.

2.3.1. The premium and default legs of a CDO. To begin with, a payment schedule $0 < t_1 < \ldots < t_n = T$ is specified, where quarter-yearly payments are a typical convention. At each payment date, the protection buyer is committed to pay the product of remaining nominal and spread (relative to the length of the preceding period) of the respective tranche. Hence, the premium leg of tranche $j$ is given by

$$PL^j = \sum_{k=1}^{n} s^j \cdot \Delta t_k \cdot \left(u^j - l^j - L^j_t\right),$$

(2.5)
where \( t_0 = 0 \) and \( \Delta t_k = t_k - t_{k-1} \). The corresponding sum of all expected discounted payments of the premium leg is then given by

\[
EDPL^j = \sum_{k=1}^{n} s^j \cdot \Delta t_k \cdot e^{-rt_k} \cdot \left( u^j - l^j - \mathbb{E}[L_t^j] \right),
\]

where a flat interest rate \( r > 0 \) is assumed and \( s^j \) is the annualized spread of this tranche. The default (or protection) leg of tranche \( j \) allows payments at any time up to maturity. A default payment in some tranche occurs if one of the companies defaults and the resulting loss affects the respective tranche. To simplify the computation, we assume that default payments are deferred to the next premium payment date. This assumption allows to conveniently express the expected discounted default leg of tranche \( j \) as

\[
EDDL^j = \sum_{k=1}^{n} e^{-rt_k} \cdot \left( \mathbb{E}[L_t^j] - \mathbb{E}[L_{t_k-1}^j] \right),
\]

where \( t_0 = 0 \). Similarly to CDS contracts, market prices for CDO tranches (with maturity \( T \)) are quoted in terms of their annualized spread \( s^j = s^j_T \). This spread is chosen such that the expected discounted payments of the default leg agree with the expected discounted payments of the premium leg of the same tranche.

### 2.3.2. The upfront payment

It has become market practice to modify the premium stream of the equity tranche using a running spread of 500 bp, which is usually below the fair spread of this first-loss piece. Therefore, a so called upfront payment is introduced which has to be paid when the contract is settled. The amount of upfront payment is quoted in percent of the nominal of the equity tranche and therefore satisfies the relation

\[
EDDL^1 = (\text{upfront}) \cdot (u^1 - l^1) + 500 \text{ bp} \cdot \sum_{k=1}^{n} \Delta t_k \cdot e^{-rt_k} \cdot \left( u^1 - l^1 - \mathbb{E}[L_t^1] \right).
\]

Depending on the terms of the contract, accrued interest for defaulted companies is often stipulated. This means that if company \( i \) defaults within the premium payment dates \( t_{k-1} \) and \( t_k \), i.e. \( \tau_i \in (t_{k-1}, t_k) \), and the total loss at time \( \tau_i \) is within tranche \( j \), then accrued interest for tranche \( j \) related to this default has to be paid at time \( t_k \) in the amount of \( s^j \cdot (\tau_i - t_{k-1}) / I \), in addition to the usual premium payment.

### 2.4. Portfolio credit default swaps

Portfolio CDS (or index CDS) are often introduced as a single tranche of a CDO covering the entire portfolio, i.e. \( l^1 = 0\% \) and \( u^1 = 100\% \). Indeed, depending on the terms of contract, this definition may hold. More precisely, the default leg compensates the insurance buyer for all losses from defaulted names in the portfolio. The premium leg is paid based on a pre-specified schedule \( 0 < t_1 < \cdots < t_n = T \) and the amount of each payment is relative to the remaining nominal of the portfolio at the time of the payment. Assuming a unit initial notional and \( I \) equally weighted assets in the portfolio with identical recovery rates \( R = R^i \), the expected discounted default leg is given by

\[
EDDL = \sum_{k=1}^{n} e^{-rt_k} \cdot \left( \mathbb{E}[L_t^k] - \mathbb{E}[L_{t_{k-1}}^k] \right),
\]

where default payments within some period are deferred to the next payment date.
Depending on the terms of contract, the remaining nominal \( N = \{ N_t \}_{t \geq 0}, N_0 = 1 \), is reduced by \( 1/\mathcal{I} \) or \( (1 - R^i)/\mathcal{I} \) after \( \tau^i \). The latter alternative corresponds to the interpretation of being a single tranche of a CDO covering the entire portfolio. Let us remark that as long as the number of defaulted companies is small, both assumptions yield very similar prices. Given the annualized spread \( s_{PCDS}^T = s_{PCDS}^T \), the expected discounted payments of the premium leg satisfy

\[
EDPL = E \left[ \sum_{k=1}^{n} s_{PCDS}^T \cdot e^{-r t_k} \cdot \Delta t_k \cdot N_{t_k} \right].
\]

Finally, the fair spread \( s_{PCDS}^{n} \) for a contract maturing in \( T \) years is obtained by equating both legs and solving for \( s_{PCDS}^{n} \).

### 2.5. \( n \)-th to default swaps

An \( n \)-th to default swap is a bilateral contract between a protection buyer and a protection seller. The payment streams of an \( n \)-th to default swap depend on the number of defaults within a reference portfolio. To make this more precise, let \( \tau(1) \leq \cdots \leq \tau(l) \) be the order statistics of the default times \( \tau^1, \ldots, \tau^l \). Initially, premium payments are payed on the schedule \( 0 < t_1 < \cdots < t_m = T \). However, these payments are conditioned to fewer than \( n \) defaults within the portfolio. The amount of each premium payment is the nominal of the contract, standardized to one unit, times the annualized spread \( s^{(n)} \), adjusted by the length of the preceding period. At the premium payment date following \( \tau^{(n)} \), i.e. when \( n \) companies defaulted, a default payment has to be paid in the amount of \( 1 - R \) times the nominal of the contract, where \( R \) is some pre-specified recovery rate. Hence, the expected discounted payments of the premium and default leg, \( EDPL^{(n)} \) and \( EDDL^{(n)} \), of an \( n \)-th to default contract are given by

\[
EDPL^{(n)} = \sum_{k=1}^{m} s^{(n)} \cdot \Delta t_k \cdot e^{-r t_k} \cdot \mathbb{P}(\tau^{(n)} > t_k),
\]

\[
EDDL^{(n)} = (1 - R) \cdot \sum_{k=1}^{m} e^{-r t_k} \cdot \mathbb{P}(t_{k-1} < \tau^{(n)} \leq t_k).
\]

Finally, the fair spread \( s^{(n)} = s^{(n)}_T \) of a contract with maturity \( T \), which allows both parties to enter the contract at zero costs, is obtained from equating both legs and solving for \( s^{(n)} \). Let us conclude this section with the remark that most \( n \)-th to default swaps are specified as first to default, larger values for \( n \) are not common in practice.

### 3. Structural-default models

All structural-default models share the interpretation of corporate default as a result of insufficient asset values. More precisely, solvency is linked to the economic fundamentals of a company via the assumption that default is triggered when the value of the firm falls below some given threshold level. Translated in probabilistic terms, the value of the considered firm is modeled as a stochastic process \( V = \{ V_t \}_{t \geq 0} \) on the filtered probability space \((\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})\). The first generation of structural-default models, e.g. [4], [59], or [5], rely on a geometric Brownian motion to model the firm-value process. Recently, this assumption was generalized using suitable Lévy processes (see, e.g. [84]). The definition of the time of default \( \tau \) is the next criterion to distinguish different models. The most simple possibility
to define $\tau$ is to only allow default at discrete dates. Alternatively, it is possible to continuously test for default, i.e. to define $\tau$ as the first-passage time of the firm-value process falling below the default threshold. This assumption seems to be more realistic, but requires the distribution of the infimum of the firm-value process for pricing applications. Considering the default threshold, there are also several options available. The most simple model is to assume a constant default threshold. This can be generalized to stochastic thresholds and to non-observable thresholds that are modeled as a random variable, see, e.g. [28] and [27]. Finally, it is possible to allow the firm’s management, seeking to maximize the firm’s equity, to endogenously choose the default threshold. This approach was proposed by [50] and [51]. In a calibration of the threshold, the company’s liabilities are often used as a proxy variable. Other popular interpretations are weighted averages of short- and long-term liabilities (see [6]) or a minimum firm value which is required to operate the company (see [5]). Finally, it is possible to consider different assumptions concerning the information of the investor. Mathematically, the most convenient assumption is to let investors continuously observe the process $V$, i.e. to let $\mathbb{F}$ be the natural filtration of $V$. This assumption is critical, as in reality the firm-value process is only revealed at the times when the balance sheet is published. These issues have been discussed by [18]. Consequently, the model of the firm-value process, the definition of $\tau$, the available information, and the default threshold fully define a structural-default model. Therefore, these assumptions implicitly specify the resulting term structure of default probabilities, based on which corporate bonds and credit derivatives are priced.

3.1. Merton’s model (1974). The first structural-default model was published by [4]. Originally, their model was designed to describe stock prices rather than the value process of a firm. Then, the observation "It is not generally realized, that corporate liabilities other than warrants may be viewed as options" (see [4], page 649) founded the idea of structural-default models. In this model, the firm-value process is a fully observable geometric Brownian motion, default is tested at maturity only, and the default threshold is constant. This idea was worked out in detail by [59], who slightly changed the underlying stochastic-differential equation (sde) of the firm-value process to include dividends and interest payments. However, the solution of this equation is still a geometric Brownian motion. More precisely, let $V = \{V_t\}_{t \geq 0}$ be defined on the filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ via

$$dV_t = V_t \cdot (\mu \, dt + \sigma dW_t), \quad V_0 > 0.$$  

This sde implies that returns of the firm-value process are decomposed into a systematic drift component and a random noise, the latter is modeled via a Brownian motion. Let us remark that the $\mathbb{P}$-drift $\mu$ of $V$ allows for different interpretations. Originally, the assumption was to choose $\mu = r$, which corresponds to the risk-neutral stock dynamics in an option pricing framework. As the firm-value process is not a traded asset, other choices are also possible. Applying Itô’s formula to the logarithm of $V_t$ yields

$$d \ln(V_t) = (\mu - 0.5 \cdot \sigma^2) \, dt + \sigma dW_t,$$

which we integrate to find the exact solution of sde (3.1), given by

$$V_t = V_0 \cdot \exp \left( (\mu - 0.5 \cdot \sigma^2) t + \sigma W_t \right).$$
To simplify the notation, let us introduce the abbreviation $\gamma = \mu - 0.5 \cdot \sigma^2$. If we assume a constant threshold level $d$, satisfying $0 < d < V_0$, and test for default at time $t$, we obtain

$$P(\tau \leq t) = \mathbb{P}(V_t \leq d) = \Phi \left( \frac{\ln(d/V_0) - \gamma \cdot t}{\sigma \cdot \sqrt{t}} \right),$$

where $\Phi(x)$ is the cumulative distribution function of a $N(0,1)$-distributed random variable.

3.1.1. Pricing in Merton’s model. Pricing a zero-coupon bond requires an explicit formula of the $t$-year default probability. As this is given in Equation (3.2), we easily obtain the following result.

**Lemma 3.1 (Pricing a defaultable zero-coupon bond in Merton’s model).** Let the interest rate $r > 0$ be constant and let $R$ denote the fixed recovery rate. If default is only considered at maturity $T$, the time zero price of a defaultable zero-coupon bond with unit face value is given by

$$P^d(0, T) = \mathbb{E} \left[ e^{-rT} \cdot 1_{\{\tau > T\}} + R \cdot e^{-rT} \cdot 1_{\{\tau \leq T\}} \right]$$

$$= e^{-rT} \cdot \mathbb{P}(\tau > T) + R \cdot e^{-rT} \cdot \mathbb{P}(\tau \leq T)$$

$$= e^{-rT} \cdot \left( 1 - \Phi \left( \frac{\ln(d/V_0) - \gamma \cdot T}{\sigma \cdot \sqrt{T}} \right) + R \cdot \Phi \left( \frac{\ln(d/V_0) - \gamma \cdot T}{\sigma \cdot \sqrt{T}} \right) \right).$$

In Merton’s model it is possible to express the recovery rate $R$ as a random variable which is revealed at the time of default. If the firm-value process is below $d$ and the companies’ liabilities are chosen as threshold level, a natural choice for the recovery rate is the fraction $V_T/d$, which corresponds to the scheme fractional recovery of firm value. With this assumption, the corresponding pricing formula can be shown to result in a non-defaultable bond plus a short position in a European put-option on the recovery rate with strike one. The second position is identical to a number of $1/d$ European put options on the firm-value process with strike $d$. Pricing coupon bonds and CDS contracts is done similarly. For this reason, we omit the respective formulas. To give an example, the premium leg of a CDS can be derived by plugging the appropriate survival probabilities from Equation (3.2) into the pricing formula. If default payments are deferred to the next premium payment date, the same is possible for the default leg.

3.1.2. Advantages and shortfalls of Merton’s model. The major advantage of Merton’s model is its simplicity and the resulting closed-form expressions for bond prices and CDS spreads. A major drawback is observed if the term-structure of bond (or CDS) spreads is compared to real curves. It turns out that the model produces spread curves that are not observed in the market. Investigating this further, the spread corresponding to the bond price of Lemma 3.1 is given by

$$S_T = -\frac{1}{T} \ln(P^d(0, T)) - r = -\frac{1}{T} \ln \left( 1 + (R - 1) \cdot \Phi \left( \frac{\ln(d/V_0) - \gamma \cdot T}{\sigma \cdot \sqrt{T}} \right) \right).$$

A term structure of spreads with $r = 0.03$, $R = 40\%$, $\sigma = 0.1$, $\gamma = r - 0.5 \cdot \sigma^2$, and $d/V_0 = 80\%$ is given in Figure 2 below. Unrealistic is the hump-shaped structure of the spread curve, which is usually not observed in the market. Using l’Hospital’s rule it can further be shown that the limit of bond and CDS spreads, as maturity decreases to zero, is zero. This is also unrealistic, as even for the shortest maturity investors ask for a risk premium.
3.2. The model of Black and Cox (1976). In [4] and [59], default is only possible at maturity. This unrealistic assumption is convenient for the computation of bond prices, as the bond-pricing formula can be rewritten in terms of the formula of a European put option. This shortcoming was corrected by [5], who criticize the original model as follows: "Furthermore, it assumes that the fortunes of the firm may cause its value to rise to an arbitrary high level or dwindle to nearly nothing without any sort of reorganization occurring in the firm’s financial arrangement. More generally, there may be both lower and upper boundaries at which the firm’s securities must take on specific values" (see [5], page 352.). To correct this unrealistic assumption, they propose to continuously test for default and to define the time of default as the first-passage time of the firm-value process below the default threshold. Formally, \( \tau \) is now defined as

\[
\tau = \inf\{t > 0 : V_t \leq d\}.
\]

This definition of default is more realistic as insufficient assets cause an immediate default. However, it requires us to compute the distribution of the running minimum of the firm-value process. Keeping the assumption of a geometric Brownian motion, this distribution is analytically tractable in the model of Black and Cox. The following result is well-known (see, e.g., [61], page 61).

Lemma 3.2 (The minimum of a Brownian motion with drift). Let \( X_t = \gamma t + \sigma W_t \) denote a diffusion over \([t_0, t_1]\), starting at \( X_{t_0} > b \), where \( b \) denotes a barrier level and define \( \Delta t = t_1 - t_0 \). We abbreviate the term

\[
P\left( \min_{t_0 \leq s \leq t_1} \{\sigma W_s + \gamma s\} > b \mid X_{t_0} \right)
\]

by \( \Phi_{b,\gamma,\sigma}^{BM}(X_{t_0}, \Delta t) \), which simplifies to

\[
1_{\{X_{t_0} > b\}} \left( \Phi \left( \frac{X_{t_0} - b + \gamma \cdot \Delta t}{\sigma \cdot \sqrt{\Delta t}} \right) - e^{-2\gamma(X_{t_0} - b)/\sigma^2} \cdot \Phi \left( \frac{-X_{t_0} + b + \gamma \cdot \Delta t}{\sigma \cdot \sqrt{\Delta t}} \right) \right).
\]

3.2.1. Pricing in the model of Black and Cox. If we continuously test for default, the pricing formula of a zero-coupon bond, assuming a fractional recovery of face value, can be written as

\[
P^d(0, T) = e^{-rT} \cdot \mathbb{P}(\tau > T) + R \cdot \int_0^T e^{-rt} d\mathbb{P}(\tau \leq t).
\]
This formula reflects the fact that a payment at the time of default must be discounted with the corresponding discount factor.

**Theorem 3.3.** Assuming a fractional recovery of face value, the price of a defaultable zero-coupon bond with maturity $T$ is given by

$$P^d(0, T) = e^{-rT} \cdot \Phi_{b, \gamma, \sigma}(0, T) + Re^{-b(\gamma - \gamma)\cdot \sigma^{-2}} \left(1 - \Phi_{b, \gamma, \sigma}(0, T)\right),$$

where $\gamma = \sqrt{\sigma^2 + 2r\sigma^2}$ and $b = \ln(d/V_0)$.

**Proof:** To evaluate the Riemann-Stieltjes integral of Formula (3.3), we make use of a lemma which is presented in [3], page 74. The claim of this lemma is that for real numbers $a$, $b$, and $c$, satisfying $b < 0$ and $c^2 > a$, we have

$$\int_{0}^{y} e^{ax} d\Phi\left(\frac{b-ex}{\sqrt{x}}\right) = \frac{d+c}{2d} \cdot h_1(y) + \frac{d-c}{2d} \cdot h_2(y),$$

with the abbreviations $d = c^2 - 2a$ and

$$h_1(y) = e^{(c-d)} \cdot \Phi\left((b-d)y^{-1/2}\right), \quad h_2(y) = e^{(c+d)} \cdot \Phi\left((b+dy)y^{-1/2}\right).$$

Using the linearity of a Riemann-Stieltjes integral in the integrator, we can evaluate the required integral using Equation (3.5). A lengthy calculation shows that the integral $\int_{0}^{T} e^{-rt} d\Phi\left(\frac{b+\gamma t}{\sigma \sqrt{t}}\right)$ simplifies to

$$\frac{1}{2} \left(1 + \frac{\gamma}{\gamma}\right) \cdot e^{b(\gamma-\gamma)a^{-2}} \cdot \Phi\left(\frac{b}{\sqrt{\gamma}}\right) + \frac{1}{2} \left(1 - \frac{\gamma}{\gamma}\right) \cdot e^{b(\gamma+\gamma)a^{-2}} \cdot \Phi\left(\frac{b}{\sqrt{\gamma}}\right).$$

Moreover, using the same result, it holds that $e^{2\gamma a^{-2}} \int_{0}^{T} e^{-rt} d\Phi\left(\frac{b+\gamma t}{\sigma \sqrt{t}}\right)$ equals

$$\frac{1}{2} \left(1 - \frac{\gamma}{\gamma}\right) \cdot e^{b(\gamma-\gamma)a^{-2}} \cdot \Phi\left(\frac{b}{\sqrt{\gamma}}\right) + \frac{1}{2} \left(1 + \frac{\gamma}{\gamma}\right) \cdot e^{b(\gamma+\gamma)a^{-2}} \cdot \Phi\left(\frac{b}{\sqrt{\gamma}}\right).$$

Combined, this yields

$$R \int_{0}^{T} e^{-rt} d\mathbb{P}(\tau \leq t) = R \cdot \left(e^{b(\gamma-\gamma)a^{-2}} \cdot \Phi\left(\frac{b}{\sqrt{\gamma}}\right) + e^{b(\gamma+\gamma)a^{-2}} \cdot \Phi\left(\frac{b}{\sqrt{\gamma}}\right)\right).$$

Some algebraic manipulations with Lemma 3.2 complete the proof. \[\square\]

Pricing CDS is closely related to the pricing of corporate bonds. Again, we are required to evaluate an expectation which depends on the distribution of $\tau$. To begin with, we use the integration by parts formula for Riemann-Stieltjes integrals and find

$$\int_{0}^{T} e^{-rt} d\mathbb{P}(\tau \leq t) = 1 - e^{-rT} \cdot \mathbb{P}(\tau > T) - r \cdot \int_{0}^{T} \mathbb{P}(\tau > t) \cdot e^{-rt} dt.$$

Since we already computed the left-hand side in closed form, the fair CDS spread is easily found from this formula after some algebraic manipulations. This yields the following theorem.

**Theorem 3.4 (Pricing CDS in the model of Black and Cox).** We consider a CDS with unit notional, continuous spread payments, and recovery rate $R$. Given
14 RUDI ZAGST AND MATTHIAS SCHERER

the constant interest rate $r > 0$, the premium that allows both parties to enter the contract at par is given by

$$s^{CDS}_T = s^{CDS} = \frac{r(1 - R)A}{1 - e^{-rT}B - A},$$

where $b = \ln(d/V_0)$, $\tilde{\gamma} = \sqrt{\gamma^2 + 2r\sigma^2}$, and

$$A = e^{b\sigma^2(\gamma-\tilde{\gamma})} (1 - \Phi_{B,\tilde{\gamma},\sigma}(0, T)), \quad B = \Phi_{B,\tilde{\gamma},\sigma}(0, T).$$

**Proof:** We use Equation (2.4) and the explicit formula of $\int_0^T e^{-rt}d\mathbb{P}(\tau \leq t)$ as computed in the proof of Theorem 3.3.

3.2.2. The limit of spreads. In Merton’s model, we already showed that it is virtually impossible for a solvent company to default within a small interval of time, due to the assumed continuity of the firm-value processes, which leads to vanishing spreads as maturity decreases to zero. In the model of Black and Cox, we show that this problem remains for bond and CDS spreads. To show this for CDS spreads, let us consider the abbreviations $A$ and $B$ from Equation (3.7) as functions of $T$.

Then, we observe that as long as the company is solvent or, in other words, as long as the distance to default $-b = \ln(V_0/d)$ is positive, we obtain the following limits

$$\lim_{T \downarrow 0} A(T) = \lim_{T \downarrow 0} A'(T) = \lim_{T \downarrow 0} B'(T) = 0, \quad \lim_{T \downarrow 0} B(T) = 1.$$

L’Hospital’s rule finally establishes

$$\lim_{T \downarrow 0} s^{CDS}_T = \lim_{T \downarrow 0} \frac{r \cdot (1 - R) \cdot A'(T)}{r \cdot e^{-rT} \cdot B(T) - e^{-rT} \cdot B'(T) - A'(T)} = 0.$$

3.3. The model of Zhou (2001). To overcome vanishing short-term spreads, [84] proposed to relax the assumption of a continuous firm-value process. He suggested to model the firm-value process as the exponential of a jump-diffusion process with normally distributed jumps. The advantage of this approach is that a jump can cause a sudden default, leading to a positive limit of short-term credit spreads. The disadvantage of this generalization is that the distribution of the minimum of a jump-diffusion process is generally not known in closed-form. Hence, numerical methods are required to derive prices for bonds and CDS in this context. In what follows, we assume that the value of a company follows a stochastic process $V = \{V_t\}_{t \geq 0}$ on the filtered probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with $V_t = V_0 \cdot \exp(X_t)$, $V_0 > 0$, where the process $X = \{X_t\}_{t \geq 0}$ is the superposition of a diffusion and a compound Poisson process, i.e.

$$X_t = \gamma t + \sigma W_t + \sum_{i=1}^{N_t} Y_i.$$

To exclude degenerated cases, we assume a non-trivial diffusion component $\sigma > 0$, a positive intensity $\lambda > 0$ of the Poisson process $N_t$, and a non-degenerated jump measure $\mathbb{P}_Y \neq \delta_0$. Again, we let $\tau$ be the first-passage time of the firm-value process below the debt level $d$. 
3.3.1. Pricing via Monte Carlo in a jump-diffusion scenario. [84] suggested
a simple algorithm for the computation of bond prices in the given framework.
First, trajectories of the firm-value process on a fine grid are sampled. Then, it
is checked at each point of the grid whether the company defaults or not. Given
this run, the resulting bond price is computed. This approach is intuitive and
straightforward to implement, but has two disadvantages. First of all, a small
discretization bias is the result of only monitoring the firm-value process on a grid.
Default probabilities are systematically underestimated, as the firm may default
between two points of the grid. If the grid is made finer, the number of samples that
have to be drawn increases accordingly, making the algorithm very slow. However,
one can significantly improve Zhou’s algorithm if the variance reduction technique
conditional Monte Carlo is used. The result is an unbiased and fast Monte Carlo
algorithm. The principal idea, see [66], of this improved algorithm is to condition
on the number of jumps, the location of the jump times, and the values of the
algorithm. The principal idea, see [66], of this improved algorithm is to condition
on the number of jumps, the location of the jump times, and the values of the
firm-value process at these times. As soon as we condition on the jumps, we know
that Brownian bridges connect the firm-value process in between. The probability
of a Brownian bridge not crossing a certain barrier \( b \) is calculated by [60]. In the
sequel we use a slightly simplified expression which is more convenient to work with.
Similar results for the maximum of a Brownian bridge can be found in [8], page 61,
or in [42], page 265.

**Lemma 3.5 (The minimum of a Brownian bridge).** Let \( X = \{ X_t \}_{t \in [t_0, t_1]} \)
denote a Brownian bridge\(^2\) over \([t_0, t_1]\) with volatility \( \sigma \), pinned at \( X_{t_0} \) and \( X_{t_1} \). Let \( b \in \mathbb{R} \)
denote an arbitrary barrier. Then, we define

\[
C_t = \{ \omega \in \Omega : \{ X_s(\omega) \}_{t_0 \leq s \leq t_1} \text{ passes } b \text{ for the first time in } [t, t + dt] \}
\]

and obtain for \( t \in (t_0, t_1) \)

\[
(3.9) \quad g(t) dt = \mathbb{P}(C_t | X_{t_0}, X_{t_1})
\]

\[
= \mathbf{1}_{\{ X_{t_0} > b \}} \cdot \frac{X_{t_0} - b}{2y \pi \sigma^2 (t - t_0)^{3/2} (t_1 - t)^{1/2}} \cdot \exp \left( -\frac{(X_{t_1} - b)^2}{2(t_1 - t)\sigma^2} - \frac{(X_{t_0} - b)^2}{2(t - t_0)\sigma^2} \right) dt,
\]

where \( y \) is defined by

\[
y = \frac{1}{\sqrt{2\pi \sigma^2(t_1 - t_0)}} \cdot \exp \left( -\frac{(X_{t_1} - X_{t_0})^2}{2\sigma^2(t_1 - t_0)} \right).
\]

By integration, we obtain the probability of \( X \) falling below the barrier \( b \)

\[
\tilde{\Phi}_{b,\sigma}^{BB}(X_{t_0}, X_{t_1}, \Delta t) = \mathbb{P} \left( \min_{t_0 \leq s \leq t_1} X_s \leq b \bigg| X_{t_0}, X_{t_1} \right)
\]

\[
= \mathbf{1}_{\{ X_{t_0} \leq b \text{ or } X_{t_1} \leq b \}} + \mathbf{1}_{\{ X_{t_0} > b \text{ and } X_{t_1} > b \}} \cdot \exp \left( -\frac{2(X_{t_0} - b)(X_{t_1} - b)}{(t_1 - t_0)\sigma^2} \right).
\]

Finally, we define \( \Phi_{b,\sigma}^{BB}(X_{t_0}, X_{t_1}, \Delta t) = 1 - \tilde{\Phi}_{b,\sigma}^{BB}(X_{t_0}, X_{t_1}, \Delta t) \) as the probability of
the Brownian bridge to remain above the threshold \( b \) within the interval \([t_0, t_1]\).

\(^2A \text{ definition and some properties of a Brownian bridge are given in [22], page 245.}\)
In what follows, we show how this result is used to price zero-coupon bonds. A similar idea leads to an efficient CDS pricing routine. The pricing formula of a zero-coupon bond, under fractional recovery of face value, can be written as the conditional expectation

\[ P^d(0,T) = \mathbb{E} \left[ \mathbb{E} \left[ 1_{\{\tau > T\}} e^{-rT} + R \cdot 1_{\{\tau \leq T\}} e^{-r\tau} \mid \mathcal{F}^* \right] \right], \]

where the \( \sigma \)-algebra

\[ \mathcal{F}^* = \sigma \{ N_T; \tau_1 < \cdots < \tau_{N_T} < T; X_{\tau_{t-}}, X_{\tau_{t}}, \ldots, X_{\tau_{t-1}}, X_{\tau_{t}}, \ldots, X_T \} \]

represents the information from the number of jumps, their location, and the values of \( X \) immediately before the jump times, at the jump times, and at maturity. With this choice of \( \mathcal{F}^* \), the outer expectation value integrates to

\[
\sum_{k=0}^{\infty} \int \int \int \mathbb{E} \left[ 1_{\{\tau > T\}} e^{-rT} + R \cdot 1_{\{\tau \leq T\}} e^{-r\tau} \right] \mid \mathcal{F}^* \cdot \prod_{j=1}^{k} \mathbb{P}_Y(dy_j) \prod_{j=1}^{k+1} \mathbb{P}_Y(dx_j) \mathbb{1}_{\{0 < \tau_1 < \cdots < \tau_k < T\}} \cdot \frac{k!}{\lambda^k} \cdot \frac{k!}{\lambda^k} e^{-\lambda T}.
\]

In this formula, \( \varphi_{\gamma \Delta \tau_j, \sigma^2 \Delta \tau_j} \) abbreviates the density function of a normal distribution with mean \( \gamma \cdot (\tau_j - \tau_{j-1}) \) and variance \( \sigma^2 \cdot (\tau_j - \tau_{j-1}) \), where \( \tau_0 = 0 \) and \( \tau_{N_T+1} = T \). The advantage of this reformulation is that we can compute the inner conditional expectation in closed form. With \( b = \ln(d/V_0) \), we obtain the following expression for \( \mathbb{E} \left[ 1_{\{\tau > T\}} e^{-rT} + R1_{\{\tau \leq T\}} e^{-r\tau} \mid \mathcal{F}^* \right] \):

\[
R \cdot \sum_{i=1}^{U} \prod_{j=1}^{i-1} \Phi_b^B(j) \int_{\tau_{i-1}}^{\tau_i} e^{-r\tau} \cdot g_i(\tau) d\tau + \prod_{j=1}^{N_T+1} \Phi_b^B(j),
\]

where \( I = \min \{ i \in \{1, \ldots, N_T\} : X_{\tau_i} \leq b \} \), \( \min \emptyset = 0 \), denotes the index of the first jump time such that \( X_{\tau_i} \) crosses the barrier and

\[
U = \begin{cases} 
I & \text{if } I \neq 0, \\
N_T + 1 & \text{if } I = 0.
\end{cases}
\]

Finally, \( \Phi_b^B(j) = \Phi_{b,0}^B(X_{\tau_{j-1}}, X_{\tau_j}, \tau_j - \tau_{j-1}) \) represents the probability of the company not defaulting within the interval \( (\tau_{j-1}, \tau_j) \) and \( g_i(t) \) is defined as in Equation (3.9), with \( X_b \) and \( X_t \) replaced by \( X_{\tau_{i-1}} \) and \( X_{\tau_i} \), respectively.

The density of the outer integral are the \( \text{Poi}(\lambda T) \) distribution of \( N_T \), conditioned on which the jump times are distributed as order statistics (see page 17 of [69]). Finally, the increments of the diffusion component in between two jumps, as well as the jump sizes, are mutually independent and distributed as \( \mathcal{N}(\gamma \Delta \tau_j, \sigma^2 \Delta \tau_j) \) and \( \mathbb{P}_Y \), respectively. The three summands of the bond price are explained as follows. As the firm-value process is only sampled at \( \tau_i \) and \( \tau_{i-1} \), we have to consider three events: default by diffusion (first summand), default by jump (second summand), and no default (third summand). The resulting payoff is then weighted with the respective conditional probability given \( \mathcal{F}^* \). A detailed proof is given in [66].
Based on the computation above, we can now evaluate the outer integral via a Monte Carlo simulation, in which we compute the inner conditional expectation in closed form. We obtain the following algorithm to estimate $P_d(0,T)$ without discretization bias.

**Algorithm 3.6 (Brownian-bridge pricing algorithm).** Choose the number of simulation runs $K$ and approximate $P_d(0,T)$ by the arithmetic mean of

$$P_1^d(0,T), \ldots, P_K^d(0,T),$$

where in the $j$-th step $P_j^d(0,T)$ is calculated as follows.

1. Draw the number of jumps $N_T$, a Poisson distribution with parameter $\lambda T$.
2. Simulate the jump times $0 < \tau_1 < \ldots < \tau_{N_T} < T$. Conditioned on $N_T$, these jump times are distributed as order statistics of $\text{Uni}(0,T)$-distributed random variables on $[0,T]$ (see, e.g. [69], page 17).
3. Generate two series of independent random variates $x_1, \ldots, x_{N_T+1}$ and $y_1, \ldots, y_{N_T}$, independent from $N_T$ with
   $x_i \sim \mathcal{N} \left( \gamma (\tau_i - \tau_{i-1}), \sigma^2 (\tau_i - \tau_{i-1}) \right), \quad y_i \sim \mathcal{P} \gamma.$
4. Calculate $X_0, X_{\tau_1}, X_{\tau_2}, \ldots, X_{\tau_{N_T}}, X_{\tau_{N_T}+1} = X_{\tau_{N_T}+1}$ by
   $$X_{\tau_0} = 0, \quad X_{\tau_i} = X_{\tau_{i-1}} + x_i, \quad \forall i \in \{1, \ldots, N_T\};$$
   $$X_{\tau_i} = X_{\tau_{i-1}} + y_i, \quad \forall i \in \{1, \ldots, N_T\}.$$
5. Determine $I, U$, and $b$ as explained above and calculate
   $$P_j^d(0,T) = \mathbb{E} \left[ 1\{\tau > T\} \cdot e^{-rT} + R \cdot 1\{\tau \leq T\} \cdot e^{-r\tau} \mid F^* \right].$$

The runtime of this algorithm depends on the expected number of jumps $\lambda T$, as the expected number of required samples and integrals that have to be calculated is $\mathcal{O}(\lambda T)$. Let us finally remark that one can easily generalize the algorithm to include stochastic recovery rates.

### 3.4. Models based on other Lévy processes.

While several complicated Lévy models have become standard for the modeling of stocks and interest rates, it is still challenging to construct a tractable firm-value model based on other Lévy processes than the Brownian motion. The reason for this is that results on the distribution of the running minimum of a Lévy process are rare. Therefore, first-passage times in a general Lévy model have to be estimated via a Monte Carlo simulation, which is usually not fast enough for an accurate calibration of the respective model. Alternatively, some specific choices of Lévy processes allow for a numerical solution of first-passage times. Examples are jump-diffusion models with two-sided exponentially distributed jumps (see, [70]), where [44] provide a formula for the Laplace transform of first-passage times and models with spectrally one-sided Lévy processes, which are also used, e.g., in [65] and [55]. Finally, it is sometimes possible to rewrite first-passage probabilities in terms of a partial-integro-differential equation, which is then solved numerically (compare [12]).
4. Intensity-based models

The limited success of continuous structural-default models in explaining credit spreads have led to the so-called intensity-based or reduced-form models. Within this class of models, the relation between default and asset value is not considered explicitly. Instead, default is modeled as a stopping time of some given hazard-rate process, i.e. the default process is specified exogenously. Hence, reduced-form models can be applied to situations where the underlying asset value is not observable, and the behavior of credit spreads for short-term maturities can be captured more realistically. The family of reduced-form models originated with \cite{37}. Since then, a long list of papers has appeared which follow this approach. Some of the most important contributions to the literature are \cite{50}, \cite{38}, \cite{47}, \cite{75}, \cite{48}, \cite{22}, and \cite{21}. Empirical evidence concerning reduced-form models is rather limited. \cite{22} finds that these models have difficulty in explaining the observed term structure of credit spreads across firms of different credit-risk qualities. \cite{16} consider specific Vasicek- and CIR-type reduced-form models and show that both models fail to account for all observed shapes of the credit spread structure. There is a heated debate which class of models - structural or reduced-form - is best (see, e.g., \cite{35}). \cite{36} compare structural and reduced-form models from an information-based perspective. They claim that the models are basically the same - the only distinction between the two model types is not whether the default time is predictable or inaccessible, but whether the information is observed by the market, as it is assumed in structural models, or not, as in reduced-form models. They conclude that for pricing and hedging, reduced-form models are the preferred methodology.

For a mathematical formulation of reduced-form models, let $T^*$ be an arbitrary but finite planning horizon and the filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ be given as in Section 2. $\mathbb{F}$ is assumed to be generated by a background process with $r$ being one of its components. For this reason, $\mathbb{F} = \{\mathcal{F}_t\}_{0 \leq t \leq T^*}$ is sometimes called the background filtration. Furthermore, let $N = \{N_t\}_{t \geq 0}$ be a counting process, i.e. a non-decreasing, integer-valued process with $N_0 = 0$, and let the default time $\tau$ be defined as the first jump of $N$, i.e. $\tau = \inf\{t > 0 : N_t > 0\}$. Furthermore, let $\mathcal{G} = \{\mathcal{G}_t\}_{0 \leq t \leq T^*}$ be the enlarged filtration $\mathcal{G} = \mathcal{F}_N \vee \mathbb{F}$, where the filtration $\mathcal{F}_N = \{\mathcal{F}_t^N\}_{0 \leq t \leq T^*}$ is generated by the counting process $N$. All filtrations are assumed to satisfy the usual conditions of completeness and right-continuity. It should be emphasized that $\tau$ is not necessarily a stopping time with respect to the filtration $\mathbb{F}$, but of course with respect to the filtration $\mathcal{G}$. If, however, the jumps of $N$ are caused by the background process, e.g. $N$ jumps whenever the background process hits a prespecified barrier, we are in the framework of the structural-default models of the previous section. We will assume throughout that for any $t \in (0, T^*)$ the $\sigma$-fields $\mathcal{F}_t$ and $\mathcal{F}_t^N$ are conditionally independent (under the pricing measure $\mathbb{P}$) given $\mathcal{F}_t$. This is equivalent to the assumption that $\mathbb{F}$ has the so-called martingale invariance property with respect to $\mathcal{G}$, i.e. any $\mathbb{F}$-martingale is also a $\mathcal{G}$-martingale (see \cite{3}, p. 167). For the technical proofs we will use another condition which is also known to be equivalent to the martingale-invariance property (see \cite{3}, p. 242): For any $t \in (0, T^*)$ and any $\mathbb{P}$-integrable $\mathcal{F}_t$-measurable random variable $X$, we have $\mathbb{E}[X|\mathcal{G}_t] = \mathbb{E}[X|\mathcal{F}_t]$. Setting $\mathcal{F}^N_{\infty} = \{\mathcal{F}_t^N\}_{t \geq 0}$, we assume that the counting process $N = \{N_t\}_{t \geq 0}$ admits a non-negative $\mathcal{F}_t$-predictable intensity $\lambda_t$ (for a detailed definition and a proof of the existence see
special cases, where in all examples the point process \( N \)

Using the martingale-invariance property, we conclude that

\[ \lambda \]

\[ \{ \] e.g. the intensity

Following a more general result of \([9]\) such that

\[ \int_0^t \lambda_s ds < \infty \text{ for all } t \geq 0 \]

\[ \int_0^\infty C_s dN_s = \int_0^\infty C_s \lambda_s ds \],

for all non-negative \( F_t \)-predictable processes \( C_t \). Following \([9]\), we then know that

\[ M = \{ M_t \}_{t \geq 0} \]

with \( M_t = N_t - \int_0^t \lambda_s ds \) is an \( F_t \)-local martingale and for any \( F_t \)-predictable process \( C_t \), such that

\[ \mathbb{E} \left[ \int_0^\infty |C_t| \lambda_s ds \right] < \infty \text{ for all } t \geq 0 \], the stochastic process defined by \( \int_0^t C_s dM_s \) is an \( F_t \)-local martingale. Thus, for small \( \Delta t > 0 \), we have

\[
\begin{align*}
\mathbb{E} \left[ N_{t+\Delta t} - N_t | G_t \right] &= 0 \cdot \mathbb{P} ( \tau \leq t | G_t ) + 1 \cdot \mathbb{P} ( t < \tau \leq t + \Delta t | G_t ) \\
&+ 0 \cdot \mathbb{P} ( \tau > t + \Delta t | G_t ) \\
&= \mathbb{P} ( t < \tau \leq t + \Delta t | G_t ).
\end{align*}
\]

Using that

\[
\lim_{\Delta t \to 0} \frac{\mathbb{E} \left[ N_{t+\Delta t} - N_t | G_t \right]}{\Delta t} = \lambda_t \cdot 1_{\{ \tau \geq t \}}, \quad \text{P-a.s.,}
\]

we get

\[
\lambda_t \cdot 1_{\{ \tau \geq t \}} = \lim_{\Delta t \to 0} \frac{\mathbb{P} ( t < \tau \leq t + \Delta t | G_t )}{\Delta t},
\]

i.e. the intensity \( \lambda_t \) can be interpreted as the arrival rate of default at time \( t \). Furthermore, the conditional default probability is given by

\[
\mathbb{P}^d ( t, T | G_t ) = \mathbb{P} ( t < \tau \leq T | G_t ) = \mathbb{E} \left[ \int_t^T \lambda_s \cdot 1_{\{ \tau \geq s \}} ds \right] G_t. 
\]

Following a more general result of \([19]\), we conclude that under some integrability conditions, the following relation holds

\[
\mathbb{P}^d ( t, T | G_t ) = 1 - \mathbb{E} \left[ e^{-\int_t^T \lambda_s ds} \right] G_t.
\]

Using the martingale-invariance property, we conclude that

\[
\mathbb{P}^d ( t, T | G_t ) = 1 - \mathbb{E} \left[ e^{-\int_t^T \lambda_s ds} \right] F_t = \mathbb{P}^d ( t, T | F_t ).
\]

The corresponding conditional survival probability \( \mathbb{P}^s ( t, T | G_t ) = \mathbb{P} ( N_T = 0 | G_t ) \) is given by

\[
\mathbb{P}^s ( t, T | G_t ) = \mathbb{P}^s ( t, T | F_t ) = \mathbb{E} \left[ e^{-\int_t^T \lambda_s ds} \right] F_t .
\]

4.1. Models with deterministic intensity. In this section we discuss some special cases, where in all examples the point process \( N \) is a Cox process, i.e. \( N = \{ N_t \}_{t \geq 0} \) is an inhomogeneous Poisson process with deterministic intensity \( \{ \lambda_t \}_{t \geq 0} \), given or conditional on the background information \( \{ F_t \}_{t \geq 0} \). Given the background information \( F_t \) and \( n \in \mathbb{N} \), this means that for \( 0 < t < T \),

\[
\mathbb{P} ( N_T - N_t = n | F_t ) = \frac{1}{n!} \cdot \left( \int_t^T \lambda_s ds \right)^n \cdot \exp \left( - \int_t^T \lambda_s ds \right).
\]

For small \( \Delta t > 0 \), this means that

\[
\mathbb{P} ( N_{t+\Delta t} - N_t = 1 | F_t ) = \lambda_t \cdot \Delta t,
\]
and, given that no jump has occurred until time $t$,

$$P^s(t, T|\mathcal{F}_t) = \mathbb{P}(N_T = 0|\mathcal{F}_t) = \exp \left( -\int_t^T \lambda_s ds \right).$$

Furthermore, the density of the time of the first jump, given the background information $\mathcal{F}_t$ and given that no jump has occurred until time $t$, is given by

$$d\mathbb{P}(\tau \leq s|\mathcal{F}_t) = \lambda_s \cdot \exp \left( -\int_t^s \lambda_u du \right) ds.$$

Let the price at time $t$ of a non-defaultable zero-coupon bond maturing in $T > t$ be given by $P(t, T)$. We can then derive the price of a defaultable zero-coupon bond as stated in the following lemma.

**Lemma 4.1 (Price of a defaultable zero-coupon bond).** Let the non-defaultable short rate $r$ and the arrival of the default be independent and the recovery rate $R$ be constant. If we assume the investor to receive the fraction $R$ at maturity (fractional recovery of treasury value), the fair price $P^d(t, T)$ of a defaultable zero-coupon bond at time $t$ with maturity $T$ satisfies, as long as $\tau > t$:

$$P^d(t, T) = \mathbb{E} \left[ e^{-\int_t^\tau r_s ds} \cdot 1_{\{N_T=0\}} \right] \mathcal{G}_t + \mathbb{E} \left[ e^{-\int_t^\tau r_s ds} \cdot R \cdot 1_{(t<\tau \leq T)} \right] \mathcal{G}_t$$

$$= P(t, T) \cdot \left( P^s(t, T|\mathcal{F}_t) + R \cdot P^d(t, T|\mathcal{F}_t) \right).$$

If $R = 0$, the zero-recovery zero-coupon bond price $P^d,\text{zero}(0, T)$ is thus given by

$$P^d,\text{zero}(t, T) = P(t, T) \cdot P^s(t, T|\mathcal{F}_t) = P(t, T) \cdot e^{-\int_0^T \lambda_s ds}.$$

For deriving the price of a CDS, let us assume that the recovery rate $R$ is known, that the default payment is immediately settled at $\tau$, and that the contract matures in $T$. Given the premium-payment schedule $0 < t_1 < \ldots < t_n = T$, the present values of the two legs in this model are given by

$$EDPL = \mathbb{E} \left[ \sum_{k=1}^n e^{-\int_{t_k}^{t_{k+1}} r_s ds} s^{CDS} \cdot \Delta t_k \cdot 1_{\{\tau > t_k\}} \right]$$

$$= \sum_{k=1}^n P^d,\text{zero}(0, t_k) \cdot s^{CDS} \cdot \Delta t_k,$$

$$EDDL = (1 - R) \cdot \mathbb{E} \left[ \int_0^T e^{-\int_0^\tau r_s ds} d\mathbb{P}(\tau \leq t) \right]$$

$$= (1 - R) \cdot \int_0^T \mathbb{E} \left[ e^{-\int_0^\tau r_s ds} \lambda_t \cdot e^{-\int_0^\tau \lambda_s ds} \right] dt$$

$$= (1 - R) \cdot \int_0^T P(0, t) \cdot \lambda_t \cdot e^{-\int_0^t \lambda_s ds} dt$$

$$= (1 - R) \cdot \int_0^T P^d,\text{zero}(0, t) \cdot \lambda_t dt,$$

where $s^{CDS} = s^{CDS}_T$ is the annualized spread of the CDS with maturity $T$, which is obtained by equating both legs and solving for $s^{CDS}$. If the interest rate and the
intensity are constant, denoted by \( r \) and \( \lambda \), respectively, we get

\[
P_{d,\text{zero}}(t, T) = e^{-(r+\lambda) \cdot (T-t)},
\]

and the present values of the premium and default leg simplify to

\[
E_{\text{DPL}} = s^{CDS} \cdot \sum_{k=1}^{n} e^{-(r+\lambda) \cdot t_k} \cdot \Delta t_k,
\]

\[
E_{\text{DDL}} = (1 - R) \cdot \int_{0}^{T} \lambda \cdot e^{-(r+\lambda) \cdot t} \, dt
\]

\[
= (1 - R) \cdot \frac{\lambda}{r+\lambda} \cdot \left( 1 - e^{-(r+\lambda) \cdot T} \right).
\]

As we have already seen, another common simplification is to assume that the payoff in the case of a default is deferred to the end of the corresponding payment period. In this case, the default leg simplifies to

\[
E_{\text{DDL}} = (1 - R) \cdot \sum_{k=1}^{n} P_{d,\text{zero}}(0, t_k) \cdot \left( e^{\int_{t_k}^{t_{k-1}} \lambda \cdot ds} - 1 \right),
\]

with \( t_0 = 0 \). Based on either simplification we can again solve for \( s^{CDS} \) and obtain the fair CDS spread of a contract maturing in \( T \).

### 4.2. Models with stochastic intensity.

In this section we assume that the default intensity is stochastic. As we assume a Cox process for \( N \), we can apply the results of the previous section by using iterated expectations. More precisely, we first solve the pricing problem conditional on a given realisation of a path of \( \lambda \). Then, we take the expectation over all possible paths for \( \lambda \). To do so, we again apply the result of [19] (see also [71], pp. 193-194) who prove that for any \( \mathcal{F}_r \)-measurable random variable \( X \) the following equality holds:

\[
E \left[ e^{-\int_{t}^{T} r \cdot ds} \cdot X \cdot 1_{\{\tau > T\}} \bigg| \mathcal{G}_t \right] = E \left[ e^{-\int_{t}^{T} r \cdot ds + \lambda \cdot ds} \cdot X \bigg| \mathcal{G}_t \right]
\]

\[
= E \left[ e^{-\int_{t}^{T} r \cdot ds + \lambda \cdot ds} \cdot X \bigg| \mathcal{F}_t \right],
\]

where we used the martingale-invariance assumption for the last equation. We thus have to calculate the expectations for the price of a zero-recovery zero-coupon bond as well as the premium and default leg of a CDS with known recovery rate \( R \), default payment immediately settled at \( \tau \), maturity in \( T \), and premium-payment schedule \( 0 < t_1 < \ldots < t_n = T \), i.e.

\[
P_{d,\text{zero}}(t, T) = E \left[ e^{-\int_{t}^{T} r \cdot ds} \cdot 1_{\{\tau > T\}} \bigg| \mathcal{F}_t \right]
\]

\[
= E \left[ e^{-\int_{t}^{T} r \cdot ds + \lambda \cdot ds} \bigg| \mathcal{F}_t \right],
\]
for the bond, and for the CDS,

\[ EDPL = \mathbb{E} \left[ \sum_{k=1}^{n} e^{-\int_{k}^{k+1} r_s \, ds} \cdot \Delta t_k \cdot I_{(\tau > t_k)} \right] \]

\[ = \sum_{k=1}^{n} P_{d, zero}(0, t_k) \cdot s^{CDS} \cdot \Delta t_k, \]

\[ EDDL = (1 - R) \cdot \mathbb{E} \left[ \int_{0}^{T} e^{-\int_{s}^{T} r_t \, dt} \, d\mathbb{P}(\tau \leq t) \right] \]

\[ = (1 - R) \cdot \int_{0}^{T} \mathbb{E} \left[ \lambda_t \cdot e^{-\int_{0}^{t} r_s \, ds + \lambda_s \, ds} \right] \, dt. \]

Even if we assume a constant non-defaultable interest rate \( r \), we now need a model for the intensity process \( \lambda = \{ \lambda_t \}_{t \geq 0} \). A famous model for the default intensity is the Vasicek model (see [80]), assuming that the intensity \( \lambda \) follows an Ornstein-Uhlenbeck process, i.e.

\[ d\lambda_t = (\theta - a \cdot \lambda_t) \, dt + \sigma \, dW_t, \quad \lambda_0 > 0, \]

with constant parameters \( \theta \geq 0, a > 0, \) and \( \sigma > 0 \), and a one-dimensional Wiener process \( W \). In this case (see, e.g. [83]), \( \lambda_t \) is normally distributed with expectation \( \exp(-a \cdot t) \cdot \lambda_0 + (1 - \exp(-a \cdot t)) \cdot \theta/a \) and variance \( (1 - \exp(-2a \cdot t)) \cdot \sigma^2/(2a) \). The survival probability can be interpreted as the price of a zero-coupon bond under the stochastic non-defaultable interest rate \( \lambda \) and is thus given by

\[ \mathbb{P}^s(t, T | \mathcal{G}_t) = \mathbb{P}^s(t, T | \mathcal{F}_t) = \mathbb{E} \left[ e^{-\int_{t}^{T} \lambda_s \, ds} \middle| \mathcal{F}_t \right] = e^{A(t, T) - B(t, T) \cdot \lambda_t}, \]

with

\[ B(t, T) = \frac{1}{a} \left( 1 - e^{-a \cdot (T - t)} \right), \]

\[ A(t, T) = \int_{t}^{T} \frac{1}{2} \sigma^2 \cdot B^2(s, T) - \theta \cdot B(s, T) \, ds \]

\[ = \left( \frac{\sigma^2}{2a^2} - \frac{\theta}{a} \right) \cdot (B(t, T) + T - t) - \frac{\sigma^2}{4a} \cdot B^2(t, T). \]

The expectation \( \mathbb{E}[\exp(-\int_{t}^{T} r_s + \lambda_s \, ds) | \mathcal{F}_t] \) can be derived analogously under specific assumptions on the process \( r \). The derivation of the expectation \( \mathbb{E}[\lambda_t \cdot \exp(-\int_{t}^{T} r_s + \lambda_s \, ds) | \mathcal{F}_t] \) is slightly more complicated. Both expectations will be derived in a more general setting within the next section.

The disadvantage of Vasicek’s model is that \( \lambda \) is not guaranteed to remain positive, due to its normal distribution, which is obviously a problem if used as default intensity. This is not possible if we instead apply the model of [14], i.e.

\[ d\lambda_t = (\theta - a \cdot \lambda_t) \, dt + \sigma \cdot \sqrt{\lambda_t} \, dW_t, \quad \lambda_0 > 0, \]

with constant parameters \( \theta \geq 0, a > 0, \) and \( \sigma > 0, \) with \( 2\theta > \sigma^2 \), and a one-dimensional Wiener process \( W \). In this case, \( \lambda_t \) follows a non-central chi-squared distribution with expectation \( \exp(-a \cdot t) \cdot \lambda_0 + (1 - \exp(-a \cdot t)) \cdot \theta/a \) and variance \( (\sigma^2/a) \cdot (\exp(-a \cdot t) - \exp(-2a \cdot t)) \cdot \lambda_0 + (\sigma^2/2a) \cdot (1 - \exp(-a \cdot t))^2 \cdot \theta/a \), given the initial value \( \lambda_0 \). For the survival probability and further examples see, e.g. [71], pp. 63-77.
5. Hybrid models

Since both structural and intensity-based models have some drawbacks, the so-called hybrid models were developed. The main characteristic of these models is given by a functional relation between the conditional probability of default and some micro- or macroeconomic factors. There are two methodologies to incorporate those factors. First, structural-default models are combined with a fundamental analysis which selects appropriate financial ratios and accounting-based measures for describing credit risk. Hence, the often criticized usage of a firm-value process as a known impact, which is unobservable in reality (see [18]), is relaxed by introducing firm-specific data as explanatory variables. Additionally, the more realistic assumption of incomplete information eliminates one of the main drawbacks of structural models, namely vanishing short-term spreads. The second and more popular approach to incorporate economic data is to link the intensity of a reduced-form model with such data. Hence, the exogenously specified default process is more tractable and the calibration of those models does not solely depend on noisy bond data. [72] introduce a so-called uncertainty index which is to be understood as an aggregation of all available information about the firm. [2] relate the credit spread to non-defaultable interest rates and a firm-specific factor. Other prominent examples are [56], [13], [15], [57], and [74]. We will briefly present the hybrid models of [72] and [2] to get an impression of the construction of these models. We will then give a more detailed analysis of the extended model of Schmid and Zagst (see [74]) and present the pricing formulas for defaultable zero-coupon bonds, default put options, and CDS in this framework. Throughout this section we assume a partial recovery of market value.

5.1. The model of Schmid and Zagst (2000). As a typical hybrid model, the Schmid and Zagst three-factor defaultable term-structure model combines elements of structural and reduced-form models. The underlying non-defaultable short rate is assumed to either follow a mean-reverting Hull-White process or a mean-reverting square-root process with time-dependent mean-reversion level. Therefore, the dynamics of the non-defaultable short rate are given by the sde

\[ dr_t = (\theta_r - a_r r_t) dt + \sigma_r r_t^\beta dW_{r,t}, \quad r_0 > 0, \]

where \( a_r, \sigma_r \) are positive constants, \( \beta \in \{0, \frac{1}{2}\} \), and \( \theta_r \) is a continuous, non-negative deterministic function. This specification implies that the short rate \( r_t \) is guaranteed to remain strictly positive if \( \beta = \frac{1}{2} \). One of the factors that determine the credit spread is the so-called uncertainty index, which can be understood as an aggregation of all information on the creditworthiness of the firm currently available. The larger the value of the uncertainty process, the lower the credit quality of the firm. The uncertainty (or signaling) process is assumed to follow a mean-reverting square-root process. The uncertainty index follows the sde

\[ du_t = (\theta_u - a_u u_t) dt + \sigma_u \sqrt{u_t} dW_{u,t}, \quad u_0 > 0, \]

where \( a_u \) and \( \sigma_u \) are positive and \( \theta_u \) is a non-negative constant. The dynamics of the short-rate spread (which is supposed to be the defaultable short rate minus the non-defaultable short rate) is given by the sde

\[ ds_t = (b_s s_t - a_s s_t) dt + \sigma_s \sqrt{s_t} dW_{s,t}, \quad s_0 > 0, \]
where $a_s$, $b_s$, and $\sigma_s$ are positive constants. Therefore, the uncertainty index has a significant impact on the mean-reversion level of the short-rate spread. The corresponding formulas for the non-defaultable and defaultable zero-coupon bonds can be found in [72].

5.2. The model of Bakshi, Madan, and Zhang (2006). In the model of [2], the dynamics of the non-defaultable short rate are described by the two-factor model

$$
\begin{align*}
dr_t &= (\omega_t - a_r r_t) dt + \sigma_r \sqrt{1 - \rho_{r,w}^2} dW_{r,t} + \sigma_r \rho_{r,w} dW_{w,t}, \quad r_0 > 0, \\
dw_t &= (\theta_w - a_w w_t) dt + \sigma_w dW_{w,t}, \quad w_0 > 0,
\end{align*}
$$

where $a_r$, $a_w$, $\sigma_r$, and $\sigma_w$ are positive constants, $\theta_w \geq 0$ and $|\rho_{r,w}| < 1$. The process $w$ is assumed to be unobservable. The short-rate spread is modeled according to

$$
\begin{align*}
ds_t &= (\Lambda_r - 1)dr_t + \Lambda_u dw_t, \quad s_0 > 0,
\end{align*}
$$

where $\Lambda_r$, $\Lambda_u$ are constants and $u$ is given by

$$
\begin{align*}
du_t &= (\theta_u - a_u u_t) dt + \frac{\sigma_u \rho_{r,u}}{\sqrt{1 - \rho_{r,w}^2}} dW_{r,t} + \sigma_u \sqrt{1 - \rho_{r,w}^2} dW_{w,t}, \quad u_0 > 0,
\end{align*}
$$

$a_u$ and $\sigma_u$ are positive constants, $\theta_u \geq 0$ and $\rho_{r,u}^2 < 1 - \rho_{r,w}^2$. Therefore, the short-rate spread is driven by a factor describing the general state of the economy and a firm-specific component. [2] use firm-specific data, such as stock prices, for $u$. The corresponding formulas for the non-defaultable and defaultable zero-coupon bonds can be found in [2] and [74]. The latter compare the models of [72], [2] and the extended model of Schmid and Zagst, which we will present in the next section.

5.3. The extended model of Schmid and Zagst (2008). The extended model of Schmid and Zagst (see [74]) is a further improvement of the original Schmid and Zagst model ([72]). It is derived from the original model by simplifying the underlying stochastic processes and by introducing a fourth factor which represents the state of the economy. The main building blocks of this approach are the non-defaultable short rate and the short-rate spread. Under the pricing measure $\mathbb{P}$, the dynamics of the non-defaultable short rate $r$ are given by a two-factor Hull-White type model:

$$
\begin{align*}
dr_t &= (\theta_{r,t} + b_r r_t - a_r r_t) dt + \sigma_r dW_{r,t}, \quad r_0 > 0, \\
dw_t &= (\theta_w - a_w w_t) dt + \sigma_w dW_{w,t}, \quad w_0 > 0,
\end{align*}
$$

where $a_r$, $\sigma_r$, $a_w$, and $\sigma_w$ are positive constants, $\theta_w$ is a non-negative constant, and $\theta_r$ is a continuous, non-negative deterministic function. The factor $w$ is supposed to be observable and exemplarily fitted to the GDP growth rate. In this way, a direct link between interest-rate levels and general economic conditions is modeled. It is assumed that increasing interest rates are driven by growing GDP rates. The dynamics of the short-rate spread evolve according to

$$
\begin{align*}
ds_t &= (\theta_s + b_{s,w} w_t - b_{s,w} s_t - a_s s_t) dt + \sigma_s dW_{s,t}, \quad s_0 > 0, \\
du_t &= (\theta_u - a_u u_t) dt + \sigma_u dW_{u,t}, \quad u_0 > 0,
\end{align*}
$$

where $a_s$, $\sigma_s$, $b_{s,w}$, $b_{s,w}$, $a_s$, $\sigma_s$ are positive constants, and $\theta_u$, $\theta_s$ are non-negative constants. Furthermore $W = (W_r, W_w, W_u, W_s)'$ is a four-dimensional standard Brownian motion on the filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$. The drift of the
short-rate spread $s$ depends on the firm-specific factor $u$ and the systematic factor $w$. Through $w$ it is ensured that non-defaultable rates and credit spreads are negatively correlated, since spreads usually widen in bear-markets and tighten in bull-markets\textsuperscript{3}. Since credit-spread levels are influenced by firm-specific risk variables (see [45]), the uncertainty index $u$ is introduced to the spread process.

5.3.1. Pricing zero-coupon bonds in the model of Schmid and Zagst. Within the given framework, we will now calculate the price of a non-defaultable zero-coupon bond in the extended Schmid and Zagst model, which is a special case of the two-factor Hull-White model (see [33]).

**Theorem 5.1 (Price of a non-defaultable zero-coupon bond).** The time $t$ price of a non-defaultable zero-coupon bond with maturity $T$ is given by

$$P(t, T) = \mathbb{E} \left[ e^{-\int_t^T r_s ds} \bigg| \mathcal{F}_t \right] = e^{A(t, T)-B(t, T)r_t-E(t, T)w_t},$$

with $B(t, T) = \frac{1}{a_r} \left(1 - e^{-a_r(T-t)}\right)$,

$$E(t, T) = \frac{b_r}{a_r} \left(1 - e^{-a_w(T-t)}\right) + \frac{e^{-a_w(T-t)} - e^{-a_r(T-t)}}{a_w - a_r},$$

$$A(t, T) = \int_t^T \frac{1}{2} \sigma_r^2 B(l, T)^2 + \frac{1}{2} \sigma_w^2 E(l, T)^2 - \theta_r(l) B(l, T) - \theta_w E(l, T) dl.$$

In analogy to Section 4, let $\mathcal{G} = \{\mathcal{G}_t\}_{0 \leq t \leq T^*}$ be the enlarged filtration $\mathcal{G} = \mathbb{H} \lor \mathbb{F}$, where the filtration $\mathbb{H}$ is generated by the default-indicator function $H_t = 1_{\{\tau \leq t\}}$, with $\tau$ denoting the default time. All filtrations are assumed to satisfy the usual conditions of completeness and right-continuity. For the sake of simplicity, we furthermore assume that $\mathcal{F}_0$ is trivial. Again, we assume that the martingale-invariance property with respect to $\mathcal{G}$ holds, i.e. any $\mathbb{F}$-martingale is also a $\mathcal{G}$-martingale. Then, we can calculate the price of a defaultable zero-coupon bond in the extended Schmid and Zagst model, which follows from an application of the Feynman-Kac Theorem, by solving the corresponding partial-differential equation (see [74]).

**Theorem 5.2 (Price of a defaultable zero-coupon bond).** The price of a defaultable zero-coupon bond at time $t < \min(\tau, T)$ is given by

$$P^{d}(t, T) = \mathbb{E} \left[ e^{-\int_t^T (r_s + s_t) dt} \bigg| \mathcal{G}_t \right] = e^{A^{d}(t, T)-B(t, T)r_t-E^{d}(t, T)w_t-C^{d}(t, T)s_t-D^{d}(t, T)u_t},$$

\textsuperscript{3}There is a negative correlation of $-0.49$ between the yearly US GDP growth rates and the S&P default probabilities of $BBB$ rated bonds for the time period from 1990 until 2002.
with \( B(t, T) = \frac{1}{a_u} (1 - e^{-a_u(T-t)}) \), \( C^d(t, T) = \frac{1}{a_u} (1 - e^{-a_u(T-t)}) \),
\[
D^d(t, T) = \frac{b_{sw}}{a_s} \left( \frac{1 - e^{-a_u(T-t)}}{a_u} + \frac{e^{-a_r(T-t)} - e^{-a_s(T-t)}}{a_u - a_s} \right),
\]
\[
E^d(t, T) = \frac{b_r}{a_r} \left( \frac{1 - e^{-a_r(T-t)}}{a_r} + \frac{(1 - e^{-a_s(T-t)}) - e^{-a_r(T-t)}}{a_r - a_s} \right) - \frac{b_{sw}}{a_s} \left( \frac{1 - e^{-a_u(T-t)}}{a_u} + \frac{e^{-a_r(T-t)} - e^{-a_s(T-t)}}{a_u - a_s} \right),
\]
\[
A^d(t, T) = \int_t^T \frac{1}{2} \left( \sigma_s^2 C^d(l, T)^2 + \sigma_u^2 D^d(l, T)^2 + \sigma_w^2 E^d(l, T)^2 + \sigma_u D^d(l, T) \right) dl - \theta_s(l) B(l, T) - \theta_u C^d(l, T) - \theta_w E^d(l, T) - \theta_u D^d(l, T) dl.
\]

[74] compare different hybrid models and found that a firm-specific factor is not able to capture all changes in credit spreads but that an additional macroeconomic factor improves the performance and helps explaining credit spreads.

5.3.2. Pricing CDS in the model of Schmid and Zagst. [73] show how to price credit default options and swaps based on the model of Schmid and Zagst. To do so, they introduce the short-rate credit spread \( s_{\text{zero}} \), describing the credit-spread process equivalent to the short-rate spread \( s \) (especially of the same quality) but with zero recovery. The zero-recovery short-rate spread is assumed to be implicitly defined by
\[
(1 - z_t) \cdot s_{\text{zero}} = s_t, \quad 0 \leq t \leq T^*,
\]
where \( s = \{s_t\}_{t \geq 0} \) is the short-rate spread process and \( z = \{z_t\}_{t \geq 0}, 0 \leq z_t < 1 \) is the recovery-rate process. Since \( s_{\text{zero}} \) allows for negative values, this setup is supposed to be a local approximation of the real-world dynamics and \( s_{\text{zero}} \) is assumed to be a sufficiently good approximation of the intensity of \( H \). If \( z_t \) is a known constant, i.e. \( z_t = z \) for all \( 0 \leq t \leq T^* \), the dynamic of the zero-recovery short-rate spread is given by
\[
ds_{\text{zero}}^t = \left( \theta_{s_{\text{zero}}} + b_{sw} u_t - b_{sw} u_t - a_{s} s_{\text{zero}}^t \right) dt + \sigma_{s_{\text{zero}}} dw_{s, t}, \quad s_{\text{zero}} > 0,
\]
where \( \theta_{s_{\text{zero}}} = \frac{\theta_s}{1 - z}, \quad b_{sw} = \frac{b_{sw}}{1 - z}, \quad b_{sw}^2 = \frac{b_{sw}}{1 - z}, \) and \( \sigma_{s_{\text{zero}}} = \frac{\sigma_s}{1 - z} \). The zero-coupon bond price with zero recovery is therefore calculated by
\[
P^{d, \text{zero}}(t, T) = e^{A^{d, \text{zero}}(t, T) - B(t, T)R_t - C^{\text{zero}}(t, T) s_{\text{zero}}^t - D^{\text{zero}}(t, T) u_t - E^{d, \text{zero}}(t, T) u_t},
\]
where \( A^{d, \text{zero}}(t, T), E^{d, \text{zero}}(t, T), C^{\text{zero}}(t, T), \) and \( D^{\text{zero}}(t, T) \) are given by the corresponding formulas for \( A^d(t, T), E^d(t, T), C(t, T), \) and \( D(t, T) \), with \( \theta_s, b_{su}, b_{sw}, \) and \( \sigma_s \) substituted by \( \theta_{s_{\text{zero}}}, b_{sw}^2, b_{sw}^2, \) and \( \sigma_{s_{\text{zero}}} \), respectively. In order to price a credit default swap, [73] first introduce the price \( V^{dp} \) of a credit-default put option. A credit-default put option is a derivative under which the beneficiary pays the guarantor a fixed up-front fee in exchange for the promise to make a fixed or variable payment in the event of a default. If the underlying reference asset of the default put is a zero-coupon bond with maturity \( T^* \), the default put has maturity \( T \), and the replacement is to the difference of par. Then, the time \( t \) price of the default put is determined by calculating the expected value
\[
V^{dp}(t, T, T^*) = \mathbb{E} \left[ \int_t^T e^{-\int_t^s r_l dl} (1 - R_u) dH_u \mid \mathcal{G}_t \right],
\]
where \( R = \{R_t\}_{t \geq 0} \) denotes the recovery process describing the payoff of the underlying bond upon default. In the case of a credit default swap with maturity \( T \leq T^* \), the up-front fee \( V^{dp}(t, T, T^*) \) is exchanged by regular payments \( s \) at pre-defined times \( t < t_1 < \ldots < t_n = T \), conditional to no prior default. Hence, the value of the credit-default put option at origination must equal the value of the credit default swap:

\[
V^{dp}(t, T, T^*) = s \cdot \sum_{k=1}^{n} P^{d,zero}(t, t_k),
\]

which is equivalent to

\[
s^{CDS} = s = \frac{V^{dp}(t, T, T^*)}{\sum_{k=1}^{n} P^{d,zero}(t, t_k)}.\]

[73] also extend this framework to coupon bonds as the respective underlyings of the credit derivatives.

6. Modeling dependence

It is a rare agreement in the scientific literature that companies do not default independent of each other. Nevertheless, an open discussion is how this dependence should be explained and modeled. It is plausible that the dependence on the same macro-economic variables, e.g. interest rates, commodity prices, or political decisions, explains correlated changes in revenues and costs of different companies, with an according influence on the default probability of these firms. Therefore, we should expect empirical default rates to be affected by these parameters. In fact, we often observe default clusters, i.e. time periods with several (or hardly any) defaults. Let us illustrate this observation by a small experiment, inspired by a comparable example in [76], given in Figure 3. The top graph of Figure 3 dis-

![S&P global default rates](image)

![Simulation with 5000 independent obligors](image)

**Figure 3.** Historical and independently simulated default rates.
plays annual historical default rates of the S&P beginning in 1981. This portfolio
contains about 5,000 companies for which we observe that empirical default rates
show time periods with several, as well as time periods with only a few defaults. In contrast, if we simulate the default of 5,000 i.i.d. obligors with the average
historical default rate as default probability, we observe that the realized default
rates are year-by-year at about the same level. Explained theoretically, the central
limit theorem makes it extremely unlikely that default rates of 5,000 independent
obligors are far from the average individual default probability.

The years prior to the recent credit crisis have shown an impressive growth
in portfolio products such as CDOs. Pricing such portfolio derivatives requires a
model of joint defaults, as these products allow the interpretation of being options
on the portfolio-loss process. Also, portfolio models are required if the regulatory
capital of financial institutions has to be assessed. Therefore, we will now address
the question on how this dependence may be modeled. A popular approach is
to start with a (well known) univariate model for each company and to introduce
dependence appropriately in a second step. We explain such concepts in the sections
below.

6.1. Dependence in structural models. In a classical structural-default
model, the quantity modeled as a stochastic process is the value process of each
firm. Therefore, it is natural to introduce default correlation by making these
stochastic processes dependent. The key problem is to find a realistic model which
is still analytically (or at least numerically) tractable. Most of today’s portfolio
models are only tractable via a Monte Carlo simulation, an exception is presented
below.

6.1.1. Vasicek’s portfolio model (1987). Vasicek’s model is the first multi-di-
mensional generalization of Merton’s structural-default model. It relies on restric-
tive assumptions on the portfolio, such as identical firms and an infinite number
of companies, to allow the use of probabilistic limiting theorems. Even if these
assumptions are unrealistic, they allow to express the portfolio-loss distribution
in closed-form, which is often more important for practical applications than less
restrictive assumptions. As in Merton’s model, each firm-value process evolves ac-
ording to a geometric Brownian motion and default is tested at discrete times only,
with \( d_i \) denoting the threshold level for firm \( i \). Dependence among the companies
is introduced via a common market factor. The advantage of such an approach is
that conditional on this market factor, all companies are mutually independent.

For each firm-value process, we initially start with independent Brownian mo-
tions \( W_{i,t} \) and assume

\[
\frac{dV_{i,t}}{V_{i,0}} = \mu_i dt + \sigma_i dW_{i,t}, \quad V_{i,0} > 0.
\]

Again, we solve this sde using Itô’s formula and rewrite \( V_{i,t} \) in terms of \( V_{i,t} \)
as

\[
V_{T,i} = V_{i,0} \cdot e^{\gamma_i \cdot (T-t) + \sigma_i \cdot (W_{T,i} - W_{i,0})} = V_{i,0} \cdot e^{\gamma_i \cdot (T-t) + \sigma_i \cdot \sqrt{T-t} \cdot X_{i,t}},
\]

where \( \gamma_i = \mu_i - 0.5 \cdot \sigma_i^2 \) and \( X_{i,t} = (W_{T,i} - W_{i,0}) / \sqrt{T-t} \) follows an \( N(0,1) \)-distribution. To incorporate correlation, we partially explain \( X_{i,t} \) by the common market factor
\( M_t \) and the idiosyncratic risk factor \( \epsilon_{i,t} \), i.e. we redefine \( X_{i,t} \) as

\[
X_{i,t} = \sqrt{\rho} \cdot M_t + \sqrt{1-\rho} \cdot \epsilon_{i,t}, \quad \rho \in (0,1),
\]
where $M_t, \epsilon^1_t, \ldots, \epsilon^I_t$ are i.i.d. $\mathcal{N}(0,1)$-distributed. Note that this construction assumes identical correlation of all firms to the market factor and it is evident that the marginal distribution of $V^i_t$ is not changed. Moreover, the correlation among two firms is now given by $\text{Cor}(X^i_t, X^k_t) = \rho$, for $i \neq k$. The next important observation is that the firms are independent conditional on the market factor $M_t$. Consequently, conditional on $M_t$, all individual default events are independent. We denote the conditional default probabilities by $p^i(M_t)$ and obtain

$$p^i(M_t) = \mathbb{P}(r^i < t| M_t) = \Phi \left( \frac{k^i_t - \sqrt{\rho} \cdot M_t}{\sqrt{1 - \rho}} \right),$$

where

$$k^i_t = \frac{\ln (d^i/V^i_t) - \gamma^i \cdot (T-t)}{\sigma^i \cdot \sqrt{T-t}}.$$

The purpose of Vasicek’s model is to introduce dependence to given companies, not to explain their individual default probabilities. Therefore, our next assumption is to consider the term structure of individual default probabilities $p_t$ as given input and we set $\Phi^{-1}(p^i_t) = k^i_t$. Applying probabilistic limiting theorems requires the assumption of identical default probabilities, identical portfolio weights, and identical recovery rates. If this holds, we have

$$p(M_t) = p^i(M_t) = \Phi \left( \frac{\Phi^{-1}(p_t) - \sqrt{\rho} \cdot M_t}{\sqrt{1 - \rho}} \right).$$

We now introduce $D_t$ as the random variable describing the fraction of defaulted companies in the portfolio up to time $t$. The distribution of $D_t$ depends on two parameters, individual default probabilities $p_t$ and the correlation $\rho$ among the companies. In what follows, we denote the distribution function of $D_t$ by $F_{p_t,\rho}(x) = \mathbb{P}(D_t \leq x)$.

The last simplification is to assume the number of companies within the portfolio to be large enough to justify the use of the strong law of large numbers to approximate $D_t$ given $M_t$ by $p(M_t)$. Hence,

$$\mathbb{P}(D_t \leq x) = \int_{-\infty}^{\infty} \mathbb{P}(D_t \leq x | M_t = y) d\mathbb{P}(M_t \leq y) \approx \int_{-\infty}^{\infty} \mathbb{1}_{p(x)}(y) d\mathbb{P}(M_t \leq y) = \Phi(y^*) = \mathbb{P}(D_t \leq x),$$

with $p(-y^*) = x$. This leads to

$$F_{p_t,\rho}(x) \approx \Phi \left( \frac{\Phi^{-1}(x) \cdot \sqrt{1-\rho} - \Phi^{-1}(p_t)}{\sqrt{\rho}} \right).$$

This approximation is continuous, strictly increasing in $x$, and maps the unit interval onto itself, so it is a distribution function, too. The influence of $\rho$ on $F_{p_t,\rho}(x)$ is illustrated in Figure 4 below. The default probability $p_t$ is set to 10% in all examples. If $\rho$ is small, hardly any default correlation is introduced. Hence, portfolio losses far from 10% are very rare. If $\rho$ increases, deviations from 10% become more likely. The explanation for this is that the probability for multiple and very few defaults increases in $\rho$. 
6.1.2. Pricing CDOs with Vasicek’s model. Pricing a tranche of a CDO requires the computation of the respective expected discounted payoffs of the premium and default legs. The fair spread is then found by equating these legs and solving this relation for \( s^j \). If we consider the CDO-pricing formula, we observe that the last remaining problem is the evaluation of \( E[L^j_t] \). In Vasicek’s model, this expectation is easily derived by numerically evaluating the integral

\[
E[L^j_t] = \int_0^1 \min \left\{ (1 - R) \cdot x, u^j \right\} - \min \left\{ (1 - R) \cdot x, l^j \right\} dF_{p_t, \rho}(x),
\]

or by rewriting this expression in terms of a bivariate normal distribution (see, [64]).

6.2. Dependence in intensity-based models. The copula approach, which we present below, is a popular concept to link marginals via a copula function. The starting point is a reduced-form model which is used to obtain individual default probabilities for all companies. Then, these marginal default probabilities are connected via a suitable copula. Let us remark at this point that other models for individual default probabilities may similarly be coupled. Copulas in two-dimensions (resp. in \( I \) dimensions) are distribution functions on the unit square (resp. \( I \)-cube) with uniform marginals. Heuristically, copulas contain all information about the dependence of two (resp. \( I \)) random variables that is required to recover the joint distribution from the marginal distributions. This intuitive interpretation is made precise in Sklar’s Theorem (see [79]), which we present for simplicity in its two-dimensional version.

**Theorem 6.1 (Sklar).** Let \( H \) be a two-dimensional distribution function with marginals \( F \) and \( G \). Then, there exists a copula \( C \) such that

\[
H(x, y) = C(F(x), G(y)), \quad \forall x, y.
\]

Uniqueness of \( C \) holds if \( H \) is continuous. Conversely, for any two distribution functions \( F, G, \) and any copula \( C, \) \( H(x, y) = C(F(x), G(y)) \) defines a two-dimensional distribution function with marginals \( F \) and \( G \).

Later, we shall see that for credit-risk modeling we need an efficient algorithm to sample a vector of Uni(0, 1)-distributed random variables with some copula \( C \).
describing their dependence. A class of copulas which is convenient to sample is the class of Archimedean copulas. The presented two-dimensional definition can be generalized to higher dimensions under some assumptions on the generator.

**Definition 6.2** (Archimedean copula). Let $\varphi : [0, 1] \mapsto [0, \infty]$ be continuous, convex, and strictly decreasing, satisfying $\varphi(0) = \infty$ and $\varphi(1) = 0$. Then,

$$C(u, v) = \varphi^{-1}[\varphi(u) + \varphi(v)], \quad \forall u, v \in [0, 1],$$

is called Archimedean copula with generator $\varphi$.

Most properties of an Archimedean copula $C$ follow directly from the definition. It is easily verified that $C$ is symmetric in its arguments, $C$ is associative, i.e. $C(C(u, v), w) = C(u, C(v, w))$, $\forall u, v, w \in [0, 1]$, $C(u, u) < u$, $\forall u \in (0, 1)$, and $(\varphi(t^\alpha))^{\beta}$ is again a generator $\forall \alpha \in (0, 1], \beta \in [1, \infty)$. However, most important for our application is that sampling from several families of Archimedean copulas is relatively simple. The following algorithm for sampling an $I$-dimensional Archimedean copula with generator $\varphi$ was published by [58]. This algorithm requires the inverse Laplace-Stieltjes transform of $\varphi^{-1}$, which we abbreviate as $\mathcal{LS}^{-1}(\varphi^{-1})$.

**Algorithm 6.3** (Marshall and Olkin).

1. Generate a random number $v$ with cumulative distribution function $F_V$, where $F_V(x) = (\mathcal{LS}^{-1}(\varphi^{-1}))(x)$ for $x \in [0, \infty)$.
2. Generate $I$ i.i.d. realizations $x_i \sim \text{Uni}(0, 1)$.
3. Compute the vector $(u_1, \ldots, u_I)$ via $u_i = \varphi^{-1}(-\ln(x_i)/v), \quad i \in \{1, \ldots, I\}$.

Algorithm 6.3 is especially efficient in large dimensions, since the extra effort for each additional dimension is not more than the computation of an additional $\text{Uni}(0, 1)$-distributed random variable. The reason for this is that dependence is introduced via the common factor $v \sim F_V$. In total, we only have to sample $I + 1$ random numbers for an $I$-dimensional observation.

In Table 2 we present examples for Archimedean copulas where $F_V(x)$ is known explicitly. For the Archimedean families of Clayton, Frank, Gumbel, and Joe, $F_V(x)$ is derived in [40]. In Table 2, $\Gamma(\alpha, \beta)$ denotes the Gamma distribution with shape parameter $\alpha \in (0, \infty)$, scale parameter $\beta \in (0, \infty)$, and density

$$\frac{\beta^\alpha \cdot x^{\alpha-1} \cdot e^{-\beta x}}{\Gamma(\alpha)}, \quad x \in [0, \infty).$$

The Stable distribution is abbreviated by $S(\alpha, \beta, \gamma, \delta; 1)$ with exponent $\alpha \in (0, 2]$, skewness parameter $\beta \in [-1, 1]$, scale parameter $\gamma \in [0, \infty)$, and location parameter $\delta \in \mathbb{R}$ (see [63]). For the families Ali-Mikhail-Haq (AMH), Frank, and Joe, $F_V$ is discrete with probability mass function $(y_k)_{k \in \mathbb{N}}$ at $k \in \mathbb{N}$.

To simulate dependent defaults, we follow [49] and [77] and assume as given the term-structure of individual default probabilities. For instance, let us assume this term structure to be defined by a reduced-form model with constant intensity, i.e.

$$1 - p_i(t) = \mathbb{P}(\tau^i > t) = e^{-\lambda_i t}, \quad \forall i = 1, \ldots, I.$$

Let us remark again that we are not restricted to reduced-form models. The reason for this choice is the term structure of default probabilities to be in closed form.
Then, we choose some $I$-dimensional copula $C$ which is easy to sample and obtain the following pricing algorithm for derivatives on credit portfolios.

**Algorithm 6.4 (Monte Carlo pricing in a copula framework).** Perform the following computations in each run of the Monte Carlo simulation. After the final run, compute the average discounted payoff of the considered portfolio derivative and take this number as estimate for the price. The required number of runs may be found from asymptotic confidence intervals which are easily computed.

1. Draw $(u_1, \ldots, u_I)$ from copula $C$. Note that for each margin $u_i \sim \text{Uni}(0,1)$.
2. Obtain the (dependent) default times via $\tau^i = p_i^{-1}(u_i)$.
3. Given $(\tau^1, \ldots, \tau^I)$, compute the portfolio-loss process $L_t$.
4. Evaluate the payoff of the considered portfolio derivative, given the realized loss process of the current run.

A generalization of Algorithm 6.4 to stochastic intensities is presented in [76]. More precisely, in a reduced-form model with stochastic intensity $\lambda^i_t$ for obligor $i$, we can proceed as follows.

**Algorithm 6.5 (Monte Carlo pricing with stochastic intensities).** In each run of the Monte Carlo simulation perform the following computations.

1. As before, draw $(u_1, \ldots, u_I)$ from copula $C$.
2. Define a grid $0 = t_0 < t_1 < \ldots < t_n = T$ on $[0, T]$ on which $\lambda^i_t$ is simulated.
3. Initialize each $\lambda^i_0$ and the default countdown processes $\gamma^i_0 = 1$.
4. Start a loop $k = 1, \ldots, n$:
   a. Simulate $\lambda^i_k$ and compute $\gamma^i_{tk} = \gamma^i_{tk-1} \cdot \exp(-\lambda^i_{tk} \cdot \Delta t_k)$.
   b. Check if $\gamma^i_{tk} \geq u_i$. If so, company $i$ defaulted within $(t_{k-1}, t_k]$. Set $\tau^i = t_k$.
5. Given $(\tau^1, \ldots, \tau^I)$, compute the portfolio-loss process $L_t$.
6. Evaluate the payoff of the considered discounted portfolio derivative of the current run.

### 6.3. Further portfolio models

The scientific literature on portfolio models, especially with a focus on the pricing of CDOs, grew exponentially over the last years. Still, most models can be interpreted as extensions of structural or intensity-based models. Concerning structural models, prominent examples are [41] and [1], who generalize the factor approach using other distributions than the normal distribution. At this point, let us mention that the two methodologies should not be seen separately of each other, as each factor model indirectly specifies a copula model. Other examples of structural models are [82] and [70], who generalize the model of [84] to a portfolio of dependent firms. Moreover, let us mention

<table>
<thead>
<tr>
<th>Family</th>
<th>$\vartheta$</th>
<th>$\varphi^{-1}(t)$</th>
<th>$F_V$</th>
</tr>
</thead>
<tbody>
<tr>
<td>AMH</td>
<td>(0, 1)</td>
<td>$(1 - \vartheta)/(e^t - \vartheta)\Gamma(1/\vartheta, 1)$</td>
<td>$y_k = (1 - \vartheta)\vartheta^{k-1}$, $k \in \mathbb{N}$</td>
</tr>
<tr>
<td>Clayton</td>
<td>(0, $\infty$)</td>
<td>$(1 + t)^{-\vartheta}$</td>
<td>$F(1/\vartheta, 1)$</td>
</tr>
<tr>
<td>Frank</td>
<td>(0, $\infty$)</td>
<td>$-\vartheta \ln(e^{-\vartheta}(e^{-\vartheta} - 1) + 1)$</td>
<td>$y_k = (1-e^{-\vartheta})^k$, $k \in \mathbb{N}$</td>
</tr>
<tr>
<td>Gumbel</td>
<td>(1, $\infty$)</td>
<td>$\exp(-t^{\vartheta})$</td>
<td>$S(1/\vartheta, 1, (\cos(\frac{\pi}{\vartheta})))^{\vartheta}, 0; 1)$</td>
</tr>
<tr>
<td>Joe</td>
<td>(1, $\infty$)</td>
<td>$1 - (1 - e^{-t})^{\vartheta}$</td>
<td>$y_k = (-1)^{k+1}(\frac{1}{k})^{\vartheta}$, $k \in \mathbb{N}$</td>
</tr>
</tbody>
</table>

**Table 2.** Parameter range, generator inverse, and corresponding $F_V$. 
[31], who propose the use of nested Archimedean copulas to model hierarchical structures such as industry sectors. Finally, a new class of models directly specifies the portfolio-loss process. Examples for these so called top-down models are [78] and [29]. The advantage of top-down models is that the portfolio-loss distribution is directly given, which often simplifies the pricing of complex portfolio derivatives.

7. Conclusion

In this article we gave an overview on the fast growing market of corporate bonds and credit derivatives. On the one hand, credit derivatives allow financial institutions to transfer default risk to third parties. On the other hand, taking this risk allows investors to achieve interesting risk-return profiles. Moreover, several hedging strategies rely on the use of credit derivatives, for instance, CDS allow to take a short position in default risk. Pricing risky bonds and credit derivatives requires a model for the time of default of the underlying firm. We presented an overview on the scientific literature on structural, intensity-based, and hybrid models.

The idea of selling a credit portfolio in slices with different seniorities, called tranches, became extremely popular in the last years. Also, several other portfolio derivatives have been introduced to the credit market, for instance n-th to default swaps. For the modeling of dependent defaults we presented two methods that are popular extensions of structural and intensity-based models. However, the latest sub-prime crises shows how difficult it is to price such correlation-driven products. Therefore, we feel that especially portfolio models have the potential for further research.

References


[79] Sklar, A. (1959), Fonctions de réparation à n dimensions et leurs marges, Publications de l'Institut de Statistique de l'Université de Paris 8, pp. 229-231.


Titles in This Series

520 Manuel E. Lladser, Robert S. Maier, Marni Mishna, and Andrew Rechnitzer, Editors, Algorithmic probability and combinatorics, 2010

519 Yves Félix, Gregory Lupton, and Samuel B. Smith, Editors, Homotopy theory of function spaces and related topics, 2010


517 Tewodros Amdeberhan, Luis A. Medina, and Victor H. Moll, Editors, Gems in experimental mathematics, 2010

516 Marlos A.G. Viana and Henry P. Wynn, Editors, Algebraic methods in statistics and probability II, 2010

515 Santiago Carrillo Menéndez and José Luis Fernández Pérez, Editors, Mathematics in finance, 2010

514 Arie Leizarowitz, Boris S. Mordukhovich, Itai Shafrir, and Alexander J. Zaslavski, Editors, Nonlinear analysis and optimization II, 2010

513 Arie Leizarowitz, Boris S. Mordukhovich, Itai Shafrir, and Alexander J. Zaslavski, Editors, Nonlinear analysis and optimization I, 2010

512 Albert Fathi, Yong-Geun Oh, and Claude Viterbo, Editors, Symplectic topology and measure preserving dynamical systems, 2010

511 Luise-Charlotte Kappe, Arturo Magidin, and Robert Fitzgerald Morse, Editors, Computational group theory and the theory of groups, II, 2010

510 Mario Bonk, Jane Gilman, Howard Masur, Yair Minsky, and Michael Wolf, Editors, In the Tradition of Ahlfors-Bers, V, 2010

509 Primitivo B. Acosta-Humáñez and Francisco Marcellán, Editors, Differential algebra, complex analysis and orthogonal polynomials, 2010

508 Martin Berz and Khodr Shamseddine, Editors, Advances in p-Adic and non-archimedean analysis, 2010

507 Jorge Arvesú, Francisco Marcellán, and Andrei Martínez-Finkelshtein, Editors, Recent trends in orthogonal polynomials and approximation theory, 2010

506 Yun Gao, Naihuan Jing, Michael Lau, and Kailash C. Misra, Editors, Quantum affine algebras, extended affine Lie algebras, and their applications, 2010

505 Patricio Cifuentes, José García-Cuerva, Gustavo Garrigós, Eugenio Hernández, José María Martell, Javier Parcet, Alberto Ruiz, Fernándo Soria, José Luis Torrea, and Ana Vargas, Editors, Harmonic analysis and partial differential equations, 2010

504 Christian Ausoni, Kathryn Hess, and Jérôme Scherer, Editors, Alpine perspectives on algebraic topology, 2009

503 Marcel de Jeu, Sergei Silvestrov, Christian Skau, and Jun Tomiyama, Editors, Operator structures and dynamical systems, 2009

502 Viviana Ene and Ezra Miller, Editors, Combinatorial Aspects of Commutative Algebra, 2009

501 Karel Dekimpe, Paul Igodt, and Alain Valette, Editors, Discrete groups and geometric structures, 2009

500 Philippe Briet, François Germinet, and Georgi Raikov, Editors, Spectral and scattering theory for quantum magnetic systems, 2009

499 Antonio Giambruno, César Polcino Milies, and Sudarshan K. Sehgal, Editors, Groups, rings and group rings, 2009

498 Nicolau C. Saldanha, Lawrence Conlon, Rémi Langevin, Takashi Tsuboi, and Pawel Walczak, Editors, Foliations, geometry and topology, 2009

497 Maarten Bergvelt, Gaywalee Yamskulna, and Wenhua Zhao, Editors, Vertex operator algebras and related areas, 2009

496 Daniel J. Bates, GianMario Besana, Sandra Di Rocco, and Charles W. Wampler, Editors, Interactions of classical and numerical algebraic geometry, 2009
495  G. L. Litvinov and S. N. Sergeev, Editors, Tropical and idempotent mathematics, 2009
494  Habib Ammari and Hyeonbae Kang, Editors, Imaging microstructures: Mathematical and computational challenges, 2009
493  Ricardo Baeza, Wai Kiu Chan, Detlev W. Hoffmann, and Rainer Schulze-Pillot, Editors, Quadratic Forms—Algebra, Arithmetic, and Geometry, 2009
492  Fernando Giráldez and Miguel A. Herrero, Editors, Mathematics, Developmental Biology and Tumour Growth, 2009
491  Carolyn S. Gordon, Juan Tirao, Jorge A. Vargas, and Joseph A. Wolf, Editors, New developments in Lie theory and geometry, 2009
490  Donald Babbitt, Vyjayanthi Chari, and Rita Fioresi, Editors, Symmetry in mathematics and physics, 2009
489  David Ginzburg, Erez Lapid, and David Soudry, Editors, Automorphic Forms and $L$-functions II. Local aspects, 2009
488  David Ginzburg, Erez Lapid, and David Soudry, Editors, Automorphic forms and $L$-functions I. Global aspects, 2009
487  Gilles Lachaud, Christophe Ritzenthaler, and Michael A. Tsfasman, Editors, Arithmetic, geometry, cryptography and coding theory, 2009
486  Frédéric Mynard and Elliott Pearl, Editors, Beyond topology, 2009
485  Idris Assani, Editor, Ergodic theory, 2009
484  Motoko Kotani, Hisashi Naito, and Tatsuya Tate, Editors, Spectral analysis in geometry and number theory, 2009
483  Vyacheslav Futorny, Victor Kac, Iryna Kashuba, and Efim Zelmanov, Editors, Algebras, representations and applications, 2009
482  Kazem Mahdavi and Deborah Koslover, Editors, Advances in quantum computation, 2009
481  Aydn Aytuna, Reinhold Meise, Tosun Terzioğlu, and Dietmar Vogt, Editors, Functional analysis and complex analysis, 2009
479  Timothy Y. Chow and Daniel C. Isaksen, Editors, Communicating mathematics, 2008
478  Zongzhu Lin and Jianpan Wang, Editors, Representation theory, 2008
477  Ignacio Luengo, Editor, Recent Trends in Cryptography, 2008
476  Carlos Villegas-Blas, Editor, Fourth summer school in analysis and mathematical physics: Topics in spectral theory and quantum mechanics, 2008
475  Jean-Paul Brasselet, José Luis Cisneros-Molina, David Massey, José Seade, and Bernard Teissier, Editors, Singularities II: Geometric and topological aspects, 2008
474  Jean-Paul Brasselet, José Luis Cisneros-Molina, David Massey, José Seade, and Bernard Teissier, Editors, Singularities I: Algebraic and analytic aspects, 2008
473  Alberto Farina and Jean-Claude Saut, Editors, Stationary and time dependent Gross-Pitaevskii equations, 2008
471  Diego Dominici and Robert S. Maier, Editors, Special functions and orthogonal polynomials, 2008

For a complete list of titles in this series, visit the AMS Bookstore at www.ams.org/bookstore/.
This volume contains survey papers on mathematical finance based on some courses given at the “Lluís Santaló” Summer School of the Real Sociedad Matemática Española, held in July 2007 at the Universidad Internacional Menéndez Pelayo, Santander (Spain). The primary topics are pathwise approximations of stochastic differential equations, Hedge funds, and credit derivatives.

The paper by L. Seco and F. Chen provides a systematic survey of hedge funds from a rigorous mathematical point of view. The related paper by M. Escobar, S. Krämer, F. Scheibl, L. Seco and R. Zagst introduces a new theoretical framework for the pricing of hedge funds’ equity, inspired by the framework of Black and Cox for the valuation of company equity as a call option.

A general framework for deriving high order, stable and tractable path-wise approximations of Stratonovich stochastic differential equations as applied to finance is the subject of the paper of L. G. Gyurkó and T. Lyons.

The paper by R. Zagst and M. Scherer is a short course on the different approaches used for pricing, hedging and risk management of credit derivatives.

Researchers and practitioners in mathematical finance will find in this book a collection of excellent, up-to-date and mathematically rigorous presentations of some of the most advanced techniques for pricing and risk management.