THREE ESSAYS IN MATHEMATICAL FINANCE

Doctoral Dissertation of:
Gaetano La Bua

Supervisor:
Prof. Daniele Marazzina

Chair of the Doctoral Program:
Prof. Irene Sabadini

Year 2016 – Cycle XXIX
Acknowledgements

First and foremost, I want to thank my advisor Daniele Marazzina. It has been an honour to be your first Ph.D. student. I would like to thank you for encouraging my research and for all your contributions of time and ideas that made my Ph.D. experience productive and stimulating.

I am very grateful to Roberto Baviera and Paolo Pellicioli for the possibility to work with them (leading to the Part 2 of this Thesis) as well as to Emilio Barucci with whom, and Daniele, I had the pleasure to conduct the research project which results in the Part 3 of this Thesis. I would particularly like to thank my PhD colleague Emanuele Nastasi for his friendship, all the chats and experience we shared that have contributed immensely to my personal and professional time at Politecnico di Milano. I also thank the other members of the Finance group at Politecnico di Milano, Carlo Sgarra and Marco Guzzetti, that have been a valuable source of good advice and collaboration.

I sincerely thank my dearest friends Maria Giovanna, Massimiliano and Gabriele for providing all the support and friendship that I needed. I would also thank all the people who contributed in some way to this challenging but extremely rewarding journey.

Finally, my sincere gratitude to my family for all their love, encouragement, and motivation. Words cannot express how thankful I am to my mother, my father and my brother as well as to my “Calvi family” for having accepted and loving me as a son.

And most of all for you, Giusy, whose unconditional love and support are more valuable than you could ever imagine. Without you, this thesis would have never been written. Thank you.
Alla mia compagna di vita, amante e amica
Semplicemente... a te
This Thesis covers three different topics in the wide area of mathematical finance. Part 1 deals with the application of matrix-variate Wishart processes in derivatives pricing. Chapter 1 is devoted to the study of existing Wishart-based Stochastic Volatility models. The introduction of this kind of process in finance is motivated by the need to describe the multidimensional structure of asset variances. There are, indeed, empirical evidences (see for example [26] and [24]) that the dynamics of the implied volatility surface is driven by several factors. This causes the standard one-factor stochastic volatility models not to be flexible enough to consistently price plain vanilla options and forward volatility sensitive derivatives (e.g. forward starting and cliquet options). It is well accepted that a multi-factor approach would be necessary to take into account the variability of the skew. Further, the pricing of derivatives written on more than a single underlying assets requires a sound modelling of the multivariate dependence structure. A large part of existing literature considered vector-valued stochastic processes to model the multidimensional stochastic evolution of asset(s) volatility. The choice of \( \mathbb{R}^d \) as state space for the volatility process, however, could lead to unsatisfactory dependence structures among variance factors. This is particularly true if we restrict ourselves to the case of affine processes. In the light of the above, it appears reasonable to consider more general multidimensional processes: recently an increasing attention has been devoted to applications of matrix-defined stochastic processes in derivatives pricing. In particular, stochastic processes defined on the cone of real positive semidefinite matrices \( \mathcal{S}_d^+ (\mathbb{R}) \) can be seen as natural candidates to model the latent volatility factors. In our analysis we focus on the so-called Wishart processes introduced in [16] as a matrix generalization of square-root processes. A remarkable feature is that the analytical tractability is fully preserved since these processes belong to the class of affine processes. Given the strict connection with the well-known CIR processes, Wishart
processes have been used to define multi-factor \cite{34} and multi-asset \cite{33} extensions of the classic Heston model. Despite the analytical tractability, the implementation of Wishart-based stochastic volatility models poses non-trivial challenges from a numerical point of view. Firstly, we discuss the models calibration and, exploiting the distributional properties of Wishart process, we propose efficient model approximations that alleviate the associated computational burden. Further we highlight the constraints that need to be satisfied in order to get a well-defined Wishart process and their impact on pricing performances. Secondly, we present simple and efficient simulation schemes for the asset price trajectories, making use of the exact sampling scheme in \cite{3}, that allow to price path-dependent derivatives.

Chapter 2 presents a new class of pricing models that extend the application of Wishart processes to the so-called Stochastic Local Volatility (or hybrid) pricing paradigm. This is a very recent approach that is meant to combine the advantages of Local and Stochastic Volatility models. Despite the growing interest on the topic, however, it seems that no particular attention has been paid to the use of multidimensional specifications for the Stochastic Volatility component. Our work tries to fill the gap: we introduce two hybrid models in which the stochastic volatility dynamics is described by means of a Wishart process. The proposed parametrizations not only preserve the desirable features of existing Wishart-based models but significantly enhance the ability of reproducing market prices of vanilla options.

Part 2, based on a research project conducted with Professor Roberto Baviera (Politecnico di Milano) and Paolo Pellicioli, is concerned with the computation of CVA (Credit Valuation Adjustment) in the presence of Wrong Way Risk. CVA can be defined as the adjustment to be made to the price of a derivative transaction in order to reflect the inherent counterparty credit risk and, since the 2007 financial crisis, it has become one of the most relevant topic in the risk management industry. Departing from the assumption of independence between derivatives exposure and counterparty default probabilities leads to an additional source of risk (Wrong Way Risk) that requires a sound modelling of the dependence structure among market risk factors and default probabilities. Hull-White approach to Wrong Way Risk in the computation of CVA is considered the most straightforward generalization of the standard Basel approach. The model is financially intuitive and it can be implemented by a slight modification of existing algorithms for CVA calculation. However, path dependency in the key quantities has non elementary consequences in the calibration of model parameters. We propose a simple and fast approach for computing these quantities via a recursion formula. In the first part of Chapter 3 we show the calibration methodology on market data and CVA computations in two relevant cases: a FX forward and an interest rate swap. In the second part of the Chapter, we show that the proposed methodology leads to a straightforward application of Hull-White model to the computation of CVA for portfolios of derivatives with
early termination features, such as American or Bermuda options. Extensive numerical results highlight the non trivial impact of early exercise on CVA.

Part 3, based on a research project conducted with Professors Emilio Barucci and Daniele Marazzina (Politecnico di Milano), is devoted to the study of the impact of relative performance based salary schemes on the risk taking incentives of asset managers. In particular, in Chapter 4 we analyze the asset management problem when the manager is remunerated through a scheme based on the performance of the fund with respect to a benchmark. We show that it is not the asymmetric- fulcrum type feature that makes the difference in preventing excessive risk taking in case of a poor performance. To prevent gambling when the performance deteriorates, it is important not to provide a fixed fee to the asset manager: remuneration should be sensitive to a very poor relative performance as in the case of a capital stake or of a management fee with flow funds. We provide empirical evidence on the mutual fund industry showing excessive risk taking in case of a very poor performance and limited risk taking in case of over-performance with respect to the benchmark. These results agree with a remuneration scheme including a fixed fee and a cap.
# Contents

## 1 On Wishart-based modelling for derivatives pricing

1 Wishart Processes in Finance: Modelling Perspectives and Numerical Techniques

1.1 Introduction ........................................ 3
1.2 Definition of Wishart process and basic properties ......................... 5
   1.2.1 Existence and uniqueness conditions .................................. 6
   1.2.2 Distribution of Wishart process and related results .................. 8
1.3 The Wishart Stochastic Volatility Model .................................. 16
   1.3.1 Connection with Heston and Bi-Heston model ......................... 18
   1.3.2 Characteristic function and calibration to market prices .......... 21
   1.3.3 New simulation schemes for the WSVM ............................... 27
1.4 The Wishart Affine Stochastic Correlation Model ......................... 32
   1.4.1 A restricted version of the model .................................... 35
   1.4.2 WASC Characteristic function ........................................ 37
   1.4.3 A new calibration procedure ....................................... 38
   1.4.4 A new simulation scheme for the WASC .............................. 43
1.5 Concluding Remarks ....................................... 47

Appendix 1.A Heston and Bi-Heston approximation of WSVM .................. 48
Appendix 1.B Jacobian matrix of WASC-Heston parameters mapping .......... 56
Appendix 1.C Calibration outputs ...................................... 58
Appendix 1.D Numerical results for new simulation schemes .................. 63

## 2 A New Class of Multidimensional Wishart-based hybrid models

2.1 Introduction ........................................ 69
2.2 The general SLV framework ..................................... 71
   2.2.1 Calibration of a SLV model ....................................... 73
## Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Calibration of LV surface</td>
<td>74</td>
</tr>
<tr>
<td>Calibration of leverage function</td>
<td>74</td>
</tr>
<tr>
<td>2.3 The Wishart Stochastic Local Volatility model</td>
<td>76</td>
</tr>
<tr>
<td>2.3.1 Model dynamics</td>
<td>76</td>
</tr>
<tr>
<td>2.3.2 Numerical Results</td>
<td>77</td>
</tr>
<tr>
<td>European options</td>
<td>78</td>
</tr>
<tr>
<td>Forward starting options</td>
<td>80</td>
</tr>
<tr>
<td>2.4 The Wishart Stochastic Local Covariance model</td>
<td>80</td>
</tr>
<tr>
<td>2.4.1 Model dynamics</td>
<td>81</td>
</tr>
<tr>
<td>2.4.2 Numerical Results</td>
<td>83</td>
</tr>
<tr>
<td>2.5 Concluding remarks</td>
<td>84</td>
</tr>
<tr>
<td>Appendix 2.A WSLVM: Numerical Results</td>
<td>87</td>
</tr>
<tr>
<td>Appendix 2.B WSLCM: Numerical Results</td>
<td>89</td>
</tr>
<tr>
<td>II On the Hull-White model for the computation of CVA with WWR</td>
<td>91</td>
</tr>
<tr>
<td>3 On a simple and effective methodology to deal with the computation of CVA with WWR within the Hull-White Model</td>
<td>93</td>
</tr>
<tr>
<td>3.1 Introduction</td>
<td>93</td>
</tr>
<tr>
<td>3.2 CVA pricing and WWR</td>
<td>95</td>
</tr>
<tr>
<td>3.3 Model description</td>
<td>97</td>
</tr>
<tr>
<td>3.4 Application to the case of linear derivatives</td>
<td>99</td>
</tr>
<tr>
<td>3.4.1 The simplest application</td>
<td>99</td>
</tr>
<tr>
<td>3.4.2 A relevant application</td>
<td>100</td>
</tr>
<tr>
<td>3.4.3 Dataset description and numerical results</td>
<td>102</td>
</tr>
<tr>
<td>3.5 The Impact of Early Exercise</td>
<td>103</td>
</tr>
<tr>
<td>3.5.1 The Pricing Problem</td>
<td>105</td>
</tr>
<tr>
<td>3.5.2 The Plain Vanilla Case</td>
<td>106</td>
</tr>
<tr>
<td>3.5.3 The Bermudan Swaption Case</td>
<td>109</td>
</tr>
<tr>
<td>3.6 Concluding Remarks</td>
<td>110</td>
</tr>
<tr>
<td>Appendix 3.A Proof of Proposition</td>
<td>113</td>
</tr>
<tr>
<td>III On the impact of relative performance on fund’s manager investment strategies</td>
<td>115</td>
</tr>
<tr>
<td>4 On relative performance, remuneration and risk taking of asset managers</td>
<td>117</td>
</tr>
<tr>
<td>4.1 Introduction</td>
<td>117</td>
</tr>
<tr>
<td>4.2 The model</td>
<td>120</td>
</tr>
<tr>
<td>4.3 Relative Performance Remuneration Schemes</td>
<td>121</td>
</tr>
<tr>
<td>4.3.1 Linear scheme</td>
<td>122</td>
</tr>
<tr>
<td>Contents</td>
<td>Page</td>
</tr>
<tr>
<td>-------------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>4.3.2 Convex scheme</td>
<td>123</td>
</tr>
<tr>
<td>4.3.3 Collar type scheme</td>
<td>124</td>
</tr>
<tr>
<td>4.3.4 Convex-concave scheme</td>
<td>126</td>
</tr>
<tr>
<td>4.3.5 Concave-convex remuneration scheme</td>
<td>129</td>
</tr>
<tr>
<td>4.4 Empirical Analysis</td>
<td>131</td>
</tr>
<tr>
<td>4.5 Conclusions</td>
<td>134</td>
</tr>
<tr>
<td>Appendix 4.A Proofs</td>
<td>136</td>
</tr>
<tr>
<td><strong>Bibliography</strong></td>
<td>145</td>
</tr>
</tbody>
</table>
Part I

On Wishart-based modelling for derivatives pricing
1.1 Introduction

Given the inherent complexity of financial markets, a wide area of research in the field of mathematical finance is devoted to developing accurate models for the pricing of contingent claims. Focusing on the stochastic volatility approach (i.e., we assume to describe asset volatility as an additional stochastic process), it appears desirable to introduce sound dynamics in order to take into account the term structure of market implied volatility and/or the presence of several assets involved in the definition of multi-asset payoffs.

In the single-asset framework, for example, the one-factor specification for the volatility dynamics seems not to be flexible enough to consistently price plain vanilla options and forward volatility sensitive derivatives (e.g., forward starting and cliquet options). It is well accepted that a multifactor approach would be necessary to take into account the variability of the skew as demonstrated, for example, in [26] and [24], where principal component analyses performed on empirical implied volatilities data show the existence of multiple explanatory factors. In such a context, the most widely used parametrization is, by far, the Heston model [63] where the stochastic volatility is described by means of a scalar mean reverting process of CIR type. An intuitive way to
generalize the Heston model is to consider multi-variate extensions of CIR processes as done in [24] and [40], respectively for the single- and multi-asset settings. In both cases the authors assumed the variance factors to be defined on $\mathbb{R}_{\geq 0}^d$ for $d > 1$. Even if straightforward, the vector-valued nature of the aforementioned approaches imposes severe limitations to the dependence structure among the variance factors in order to keep analytical tractability: affinity of the overall model requires the scalar CIR processes to be independent. A wider generalization of square-root processes, and of CIR ones in particular, is given by the so-called Wishart processes introduced in [16]. These processes belong to the class of affine diffusions defined on the cone of positive semi-definite matrices extensively studied in [29]. Recently, Wishart processes have been applied in finance: the pioneering approach in [55] makes use of a Wishart process to describe the stochastic dynamics of the variance-covariance matrix among various assets. The idea is further extended in [33] by introducing the Wishart Affine Stochastic Correlation model (WASC) and adapted to the single-asset case in [34] where the Wishart Stochastic Volatility model (WSVM) is defined as a matrix-variate generalization of the Heston model. A remarkable feature of both WSVM and WASC is that they can account for non-trivial dependency among variance factors and asset(s) returns still preserving the analytical tractability. The resulting models, indeed, belong to the class of affine models in the spirit of [41].

Despite the analytical tractability, however, the implementation of this class of models is not straightforward. Firstly, as shown in [32], model calibration to market data turns out to be a delicate task due to the high dimensionality of the optimization problem to deal with. The issue is worsened by the fact that characteristic functions of log-asset prices, even in known in closed formula, are computationally intensive since they involve the evaluation of functions of matrix argument. As a consequence, standard calibration routines - that rely on transform-based techniques to price vanilla options, are proven to be inadequate in realistic applications. Unfortunately, the existing approximations available in literature seem to deal just with restrictive parameters settings [49] or provide reliable results only for options with moneyness close to the at-the-money level [32]. By exploiting the distributional properties of Wishart process, we construct efficient model approximations that mitigate the complexity of the calibration problem by replacing the Wishart-based characteristic function with those of simpler affine models. In the single-asset case, we show that, for an appropriate choice of parameters, both Heston [63] and the Bi-Heston [24] models may provide a reliable approximation of WSVM. For the multi-asset case, we make use of the distributional law of diagonal elements of Wishart process to connect the WASC calibration problem to the Heston one. In both cases, we provide the analytical form of the gradient of calibration problem objective function with respect to Wishart-based parameters allowing for a further reduction in the computational burden. This methodology extends the remarkable result provided in [30] where the gradient of a European call with respect to Heston model pa-
rameters is derived explicitly. Additionally, a thorough analysis of calibration outputs is carried out. The results, in line with other evidences in literature, show that optimal parameters do not satisfy the conditions to have a well defined Wishart process. Enforcing such conditions could have a significant impact on the pricing performance of these models. It seems, however, that the topic has not been developed in literature so far. For the first time, then, we study the effects produced by such constraints on the model induced implied volatility surfaces.

Secondly, simulation of price trajectories for this class of models is still an open problem: as far as we know, no specific algorithms have been introduced to deal with the discretization of asset dynamics within the WASC model, while the schemes proposed for the WSVM seem to face a severe trade-off in terms of consistency and computational complexity. For example in [50] several schemes are proposed as generalization of already existing schemes for 1-factor models. While simple and easy to implement, they rely on parameters assumptions hardly met in realistic market conditions. An exact simulation scheme for the WSVM is proposed in [70]: authors exploit the remarkable exact sampling method for the Wishart process devised in [3]. Despite its accuracy, the approach is rather involved and time consuming since it requires the computation of matrix-valued special functions to sample from the conditional distribution of the log-price given the realization of the Wishart process. For both models we propose ad hoc simulation schemes that embed the exact sampling scheme in [3] and turn out to be accurate and simple to implement. Not only these schemes allow to efficiently price path-dependent options in the existing models but they will be extremely useful in implementing the new pricing models presented in Chapter 2.

This Chapter is organized as follows. Section 1.2 introduces the Wishart process and presents its main properties. Section 1.3 deals with the analysis of WSVM and illustrates the original contributions in terms of model approximation and simulation schemes devised for the single-asset case. Section 1.4 delivers an analogous study of WASC. Finally, Section 1.5 contains some concluding remarks.

1.2 Definition of Wishart process and basic properties

Given the importance that Wishart process plays in the proposed framework, we start our work with a brief review of its main properties.

**Definition 1 (Wishart process).** Let $W(t)$ be a $d \times d$ Brownian motion (i.e. a matrix of $d \times d$ independent scalar Brownian motions) and $\mathbb{S}_d^+$ the set of real $d \times d$ positive semidefinite matrices. We define the Wishart process as the solution on $\mathbb{S}_d^+$ of the following SDE:

$$d\Sigma(t) = (\Omega\Omega^T + M\Sigma(t) + \Sigma(t)M^T)dt + \sqrt{\Sigma(t)} dW(t) Q + Q^T dW^T(t) \sqrt{\Sigma(t)},$$

(1.1)
Chapter 1. Wishart Processes in Finance

\[ \Sigma(0) = \Sigma_0 \in S_d^+ (\mathbb{R}) \]

with \( \Omega, Q, M \in \mathcal{M}_d(\mathbb{R}) \) (the set of real \( d \times d \) square matrices). \[ 1 \]

In the most general framework we can think of each element of all matrices appearing in (1.1) to be non null. In particular, in order to embed mean-reversion and stationarity, we consider matrix \( M \) to have only eigenvalues with negative real part. Furthermore, we relate the deterministic part of the drift in (1.1), \( \Omega \Omega^\top \), to the expected long-term value of the process, denoted with \( \Sigma_\infty \), by means of the equation

\[ - \Omega \Omega^\top = M \Sigma_\infty + \Sigma_\infty M^\top. \] (1.2)

From (1.1) and (1.2) we can easily sketch the close connection existing between Wishart and CIR processes. Indeed, Wishart processes have been firstly introduced in \([16]\) as a multidimensional extension of classic square root process: not surprisingly if we set \( d = 1 \) in (1.1), we end up with a scalar CIR process defined by the SDE

\[ dv(t) = \kappa(\theta - v(t))dt + \eta \sqrt{v(t)}dw_v(t), \quad v(0) = v_0, \] (1.3)

with \( \kappa, \theta, \) and \( \eta \) strictly positive parameters, \( v_0 \geq 0 \) and \( w_v(t) \) a scalar Brownian motion.

A remarkable feature of Wishart process is that it entails a non-trivial dependence structure among its elements. Indeed, it holds that

\[ d [\Sigma_{ij}(t), \Sigma_{kl}(t)] = (\Sigma_{ik}(t)Q^*_jl + \Sigma_{il}(t)Q^*_jk + \Sigma_{jk}(t)Q^*_il + \Sigma_{jl}(t)Q^*_ik) dt, \] (1.4)

where the notation \([\cdot, \cdot]\) refers as usual to the quadratic covariation of two stochastic processes and \( Q^* = Q^\top Q \). As we will see in the following, this property constitutes an unicuum among the affine generalizations of (1.3).

1.2.1 Existence and uniqueness conditions

Wishart processes are a particular case of affine processes defined on the cone \( S_d^+ (\mathbb{R}) \) for which general results about existence and uniqueness of solutions are provided in \([29]\) and \([81]\). We report the main result (as formulated in \([3]\)) since it plays a crucial role in our framework.

**Proposition 1** (Affine models in \( S_d^+ (\mathbb{R}) \)). Let \( X(t) \) be a generic affine process with continuous trajectories defined in \( S_d^+ (\mathbb{R}) \) by the following SDE

\[ X(t) = X(0) + \int_0^t (D_X + \mathcal{L}[X(s)]) ds + \int_0^t \left( \sqrt{X(s)} dW(s) C_X + C^\top_X dW^\top(s) \sqrt{X(s)} \right), \] (1.5)

where \( X(0), D_X \in S_d^+ (\mathbb{R}), C_X \in \mathcal{M}_d(\mathbb{R}), \mathcal{L} : S_d^+ (\mathbb{R}) \to S_d^+ (\mathbb{R}) \) is a linear transformation. Such process admits a unique weak solution in \( S_d^+ (\mathbb{R}) \) if

\[ ^1 \text{We further use the notation } \Sigma_{i,j}(t) \text{ to indicate the element in the } i \text{-th row and } j \text{-th column of } \Sigma(t). \text{ If the time dependence is omitted, we assume to refer to the elements of } \Sigma_0. \]
1.2. Definition of Wishart process and basic properties

a) \( D_X - (d-1)C_X^T C_X \in S_d^+ (\mathbb{R}) \).

b) \( \forall P_1, P_2 \in S_d^+ (\mathbb{R}) \) s.t. \( \text{Tr} [P_1 P_2] = 0 \Rightarrow \text{Tr} [\mathcal{L}(P_1) P_2] \geq 0 \), where \( \text{Tr} [\cdot] \) is the trace of a square matrix (i.e., the sum of the elements on the main diagonal).

If \( X(0) \) is in the set of real positive definite matrices \( S_d^{++} (\mathbb{R}) \) and condition a) is replaced by the stronger requirement

c) \( D_X - (d+1)C_X^T C_X \in S_d^+ (\mathbb{R}) \),

then there exist a unique strong solution to (1.5) in \( S_d^{++} (\mathbb{R}) \).

Proof. See [29] or [81].

Corollary 1 (Wishart process). A direct comparison shows that we can get the Wishart SDE (1.1) from (1.5) by setting

- \( D_X = \Omega \Omega^\top \);
- \( C_X = Q \);
- \( \mathcal{L} [P_0] = MP_0 + P_0 M^\top \).

Moreover, if we assume a more restrictive parametrization for the deterministic part of the drift

\[
\Omega \Omega^\top = \beta Q^\top Q,
\]

(1.6)

conditions a) and c) of Proposition[7] are satisfied as soon as

\[
\beta \geq d - 1,
\]

(1.7)

\[
\beta \geq d + 1,
\]

(1.8)

respectively, where the real positive parameter \( \beta \) plays the role of Feller’s condition in the univariate case. Additionally if condition a) is not met the whole process is not well defined.

For the rest of the paper we consider a Wishart process defined by (1.1) and (1.6) as usually done in financial literature.

A significant constraint has thus to be imposed on parameter \( \beta \). In the case \( d = 2 \), for example, we must require \( \beta \geq 1 \). As we are going to see, this condition is not usually met when we perform a straight calibration of Wishart-based pricing models to market prices of plain vanilla options.
1.2.2 Distribution of Wishart process and related results

In this section we present more insights about the analogy between Wishart and CIR processes by formalizing some distributional features of the Wishart process. Exploiting the affine nature of the Wishart process, we have that its characteristic function is an exponential affine transformation of the initial state as shown in the following Proposition:

**Proposition 2 (Characteristic function of Wishart process).** Let \( \Lambda \) be a real symmetric \( d \times d \) matrix, \( t \geq 0 \) and \( T - t = \tau > 0 \). The (conditional) characteristic function of the Wishart process defined by (1.1) and (1.6) is

\[
\phi_{\Sigma}(\Lambda, \tau) = \mathbb{E} \left[ \exp \left( \imath \text{Tr} \left[ \Lambda \Sigma(T) \right] \right) \mid \Sigma(t) \right] = \exp \left( \text{Tr} \left[ A_{\Sigma}(\Lambda, \tau) \Sigma(t) \right] + b_{\Sigma}(\Lambda, \tau) \right)
\]

where \( \imath \) is the imaginary unit (i.e. \( \imath = \sqrt{-1} \)) and matrix \( A_{\Sigma}(\Lambda, \tau) \) and scalar function \( b_{\Sigma}(\Lambda, \tau) \) are such that

\[
\text{Tr} \left[ A_{\Sigma}(\Lambda, \tau) \Sigma(t) \right] = \text{Tr} \left[ \imath \Lambda (\mathbb{I}_d - 2\imath \Theta(\tau) \Lambda)^{-1} \Gamma(\tau) \right]
\]

\[
b_{\Sigma}(\Lambda, \tau) = -\frac{\beta}{2} \text{Tr} \left[ \log \left( \left( \mathbb{I}_d - 2\imath \Theta(\tau) \Lambda \right) \exp \left( \tau M^\top \right) \right) - \tau M \right].
\]

The additional matrix functions appearing in the above equations are given by

\[
\Gamma(\tau) = \exp(\tau M) \Sigma(t) \exp(\tau M^\top)
\]

\[
\Theta(\tau) = \int_0^\tau \exp(uM) Q^\top Q \exp(uM^\top) \, du.
\]

**Proof.** Due to affinity, it is possible to show that \( A_{\Sigma}(\Lambda, \tau) \) is solution of a matrix Riccati ODE and \( b_{\Sigma}(\Lambda, \tau) \) is obtained by direct integration. We refer to [56] for a general discussion of the solving technique and [3] for the explicit derivation of the result. Here we simply prove that the matrix \( G = (\mathbb{I}_d - 2\imath \Theta(\tau) \Lambda) \) is invertible, i.e. \( \det[G] \neq 0 \). Matrix \( \Theta(\tau) \) is trivially positive semi-definite (since it is the integral of a positive semi-definite matrix) and let \( \mathcal{H} = \Theta(\tau) \). We then have

\[
\det[G] = \det[\mathbb{I}_d - 2\imath \mathcal{H}^\top \Lambda] = \det[\mathbb{I}_d - 2\imath \mathcal{H}^\top \mathcal{H}] = \det[\mathbb{I}_d - 2\imath \mathcal{D}] = \det[\mathbb{I}_d - 2\imath \mathcal{D}],
\]

where the second equality is a direct application of the Sylvester’s determinant identity and the third follows from the spectral theorem with \( \mathcal{D} \) the diagonal matrix whose entries are the eigenvalues of the symmetric real matrix \( \mathcal{H}^\top \mathcal{H} = \mathcal{P} \mathcal{D} \mathcal{P}^\top \). By recognizing that \( \det[\mathbb{I}_d - 2\imath \mathcal{D}] \) is the characteristic polynomial of the diagonal imaginary matrix \( 2\imath \mathcal{D} \) evaluated at 1 we immediately get the invertibility of \( G \). 

\[96\text{Furthermore, in the following we assume it to be non singular.}\]
1.2. Definition of Wishart process and basic properties

As a consequence of the analytical tractability of Wishart process and of the knowledge of its characteristic function, we are able to present two additional results regarding the distribution of diagonal elements and trace of the Wishart process.

**Corollary 2** (Distribution of elements on the main diagonal of Wishart process). Let \( \Sigma_i(t) = \Sigma_{i,i}(t) \) be the \( i \)-th element on the main diagonal of \( \Sigma(t) \) and \( F_{\chi^2}(x; \nu, \delta) \) the cumulative distribution function of a non-central chi-square random variable with \( \nu \) degrees of freedom and non-centrality parameter \( \delta \). Then, for a fixed \( T > t \),

\[
\Pr \left[ \Sigma_i(T) \leq \nu \mid \Sigma(t) \right] = F_{\chi^2} \left( \frac{\nu}{\theta_i}, \beta, \delta_i \right) \tag{1.12}
\]

where \( \delta_i = \frac{\gamma_i}{\vartheta_i} \) with \( \Gamma(\tau) = (\gamma_{i,j})_{1 \leq i,j \leq d} \) and \( \Theta(\tau) = (\vartheta_{i,j})_{1 \leq i,j \leq d} \). For ease of notation, we also set \( \gamma_i = \gamma_{i,i} \) and \( \vartheta_i = \vartheta_{i,i} \).

**Proof.** We use the fact that \( \exp \left( \text{tr} \left[ \log(G) \right] \right) = \det [G] \) for any matrix invertible matrix \( G \) to write

\[
\exp (b_\Sigma(\Lambda, \tau)) = \det \left[ \left( I_d - 2t \Theta(\tau) \Lambda \right) \exp \left( \tau M^\top \right) \right]^{-\frac{\beta}{2}} \exp \left( \frac{\beta}{2} \text{tr} [M] \tau \right)
\]

and (1.9) becomes

\[
\phi_\Sigma(\Lambda, \tau) = \det \left[ I_d - 2t \Theta(\tau) \Lambda \right]^{-\frac{\beta}{2}} \exp \left( \text{tr} \left[ t \Lambda \left( I_d - 2t \Theta(\tau) \Lambda \right)^{-1} \Gamma(\tau) \right] \right). \tag{1.13}
\]

Let \( \lambda \) be a real variable and \( e^{d}_{i,i} = (1)_{k=\ell=i} \) \( 1 \leq k, \ell \leq d \), then by setting \( \Lambda_i = \lambda e^{d}_{i,i} \), the characteristic function of \( \Sigma_i(T) \) is

\[
\phi_{\Sigma_i}(\lambda, \tau) = \phi_\Sigma(\Lambda_i, \tau) = \left( 1 - 2t \lambda \vartheta_i \right)^{-\frac{\beta}{2}} \exp \left( \frac{t \lambda \gamma_i}{1 - 2t \lambda \vartheta_i} \right) \tag{1.14}
\]

from which (1.12) follows by the definition of non-central chi-square distribution.

An alternative way to obtain (1.12) is to recognize that (1.13) is the characteristic function associated to the non-central Wishart distribution with scale matrix \( \Theta(\tau) \) and non-centrality matrix \( \Gamma(\tau) \) and apply results in [75]. For the sake of completeness, we point out that an analogous claim is shown in [32]. Our formulation, however, gives a direct interpretation of parameters involved in the distribution of \( \Sigma_i(T) \) in terms of matrices describing the Wishart process. As we will see, this turns out to be particularly useful for computational purposes.

An important consequence of (1.12) is that we can define an exact mapping between \( \Sigma_i(T) \) and a CIR process:

---

Here we exploit the invertibility of \( (I_d - 2t \Theta(\tau) \Lambda) \exp \left( \tau M^\top \right) \) which is given by the invertibility of both \( G \) (proved in the previous Proposition) and \( \exp \left( \tau M^\top \right) \) (by definition of matrix exponential).
Chapter 1. Wishart Processes in Finance

**Proposition 3 (CIR process mapping \( \Sigma_i(T) \)).** Let \( v(t) \) be a CIR process defined by (1.3). For a fixed \( T > t \), it holds that (conditionally on \( v(t) \) and \( \Sigma(t) \) respectively) \( v(T) \) and \( \Sigma_i(T) \) share the same distribution provided that

\[
v(t) = \Sigma_i(t), \quad \kappa = -\frac{1}{t} \log \left( \frac{\gamma_i}{v(t)} \right), \quad \eta = 2 \sqrt{\frac{\vartheta_i \kappa}{(1 - e^{-\kappa t})}}, \quad \theta = \frac{\beta \eta^2}{4 \kappa},
\]

where \( \gamma_i \) and \( \vartheta_i \) have been introduced in Corollary 2.

**Proof.** The correspondence of the distributions relies on the properties of the CIR process. We refer to [28] for the details.

We now focus our attention on the Wishart process trace, i.e. the process \( V(t) = \text{Tr} [\Sigma(t)] \). In [34] it is shown that via direct application of Itô’s lemma and some algebra the dynamics of \( V(t) \) is given by

\[
dV(t) = \left( \text{Tr} \left[ \beta Q^\top Q \right] + 2 \text{Tr} [M \Sigma(t)] \right) dt + 2 \text{Tr} \left[ \sqrt{\Sigma(t)} dW(t) Q \right]. \tag{1.19}
\]

The following result, firstly appeared in [75] for a generic non-central Wishart distribution, allows to express the trace of Wishart process as a weighted sum of \( d \) independent non-central chi-square random variables:

**Corollary 3 (Distribution of the Trace of Wishart process).** Let \( V(t) \) be the sum of the elements of the main diagonal of \( \Sigma(t) \) and \( \chi^2_\nu(\delta) \) be a non-central chi-square random variable with \( \nu \) degrees of freedom and non-centrality parameter \( \delta \). For a fixed \( T > t \), let \( Q \) be the orthogonal matrix that diagonalizes \( \Theta(\tau) \), that is \( Q^\top \Theta(\tau) Q = \mathcal{E} = \text{diag} [\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_d] \). Then \( V(T) \) admits the representation

\[
V(T) = \sum_{i=1}^{d} \varepsilon_i \chi^2_\nu \left( \frac{\zeta_i}{\varepsilon_i} \right) \tag{1.20}
\]

where \( \zeta_i \) is the \( i \)-th diagonal element of \( Q^\top \Gamma(\tau) Q \).

**Proof.** We propose a new simple proof of the result by exploiting (1.13). Let \( \lambda \) be a real variable (different from the one in (1.14)), then by setting \( \Lambda_V = \lambda I_d \), the characteristic function of \( V(T) \) is

\[
\phi_V(\lambda, \tau) = \phi_\Sigma(\Lambda_V, \tau) = \det \left[ I_d - 2i \lambda \Theta(\tau) \right]^{-\frac{d}{2}} \exp \left( i \lambda \text{Tr} \left[ (I_d - 2i \lambda \Theta(\tau))^{-1} \Gamma(\tau) \right] \right)
\]
1.2. Definition of Wishart process and basic properties

\[ \det \left[ I_d - 2t\lambda QE Q^T \right]^{-\frac{d}{2}} \exp \left( t\lambda \text{Tr} \left[ (I_d - 2t\lambda QE Q^T)^{-1} \Gamma(\tau) \right] \right) = \det \left[ I_d - 2t\lambda E \right]^{-\frac{d}{2}} \exp \left( t\lambda \text{Tr} \left[ Q^T \Gamma(\tau) Q (I_d - 2t\lambda E)^{-1} \right] \right) = \prod_{i=1}^{d} (1 - 2t\lambda \varepsilon_i)^{-\frac{d}{2}} \exp \left( \frac{t\lambda \zeta_i}{1 - 2t\lambda \varepsilon_i} \right) \] (1.21)

where the last equality comes from the properties of determinant and trace.

From the knowledge of the characteristic function, we can also easily compute the (conditional) moments of \( V_T \):

\[ \mathbb{E} [ V(T) | \Sigma(t) ] = \text{Tr} [ \Gamma(\tau) + \beta \Theta(\tau) ] \] (1.22)

\[ \text{Var} [ V(T) | \Sigma(t) ] = 2 \text{Tr} [ (2\Gamma(\tau) + \beta \Theta(\tau)) \Theta(\tau) ] \] (1.23)

and so on, with higher moments that get more and more involved.

The representation (1.20) allows to consider \( V(T) \) in terms of non-negative definite quadratic forms in non-central normal variables (see for example and [69] and [74]) typically arising in statistical applications. As we will see in the next section, it could be convenient to dispose of some approximations for the distribution of \( V(T) \). A feasible tool is proposed in [78] where the distribution of a non-negative quadratic form is approximated by means of an affine transformation of a non-central chi-square random variable. The unknown parameters are then chosen so that the first three cumulants of the quadratic form are matched and the difference in kurtosis is minimized. In our case this means that

\[ \Pr [ V(T) > \varrho | \Sigma(t) ] \approx \Pr \left[ \alpha_1 \lambda_{\nu^*}(\delta^*) + \alpha_0 > \varrho \right] \] (1.24)

where \( \nu^* \) and \( \delta^* \) are, respectively, the optimal degrees of freedom and non-centrality parameters while \( \alpha_1 = \sqrt{\frac{\text{Var}[V(T)\Sigma(t)]}{2(\nu^* + 2\delta^*)}} \) and \( \alpha_0 = \mathbb{E} [ V(T) | \Sigma(t) ] - \alpha_1 (\nu^* + \delta^*) \). Even if appealing from a theoretical point of view, the approximation (1.24) turns out to be inadequate in those cases in which \( V(t) \) describes the instantaneous variance of an asset returns. We here develop an alternative approximation inspired by the idea to exploit a non-central chi-square random variable to approximate \( V(T) \). Instead of considering an affine transformation, however, we just use a scaled non-central chi-square random variable with parameters fitted to match (1.22) and (1.23). In other words, we aim at approximating \( V \) with a scalar CIR process as given in the following Proposition:

**Proposition 4 (Moment-matching CIR process approximation of \( V(T) \)).** Let \( v(t) \) be a CIR process defined by (1.3). For a fixed \( T > t \), \( v(t) \) is an 2-moment matching

---

4Let \( X \) be a random variable with characteristic function function \( \phi_X(\lambda) = \mathbb{E} [\exp(i\lambda X)] \), the computation relies on the standard formula \( \mathbb{E} [X^n] = i^{-n} \phi_X^{(n)}(0) \). Additionally, we exploit the fact that for a matrix \( U = U(\lambda) \) it holds that \( \frac{d}{d\lambda} \det[U] = \det[U] \text{Tr} \left[ U^{-1} \frac{dU}{d\lambda} \right] \) and \( \frac{dU^{-1}}{d\lambda} = -U^{-1} \frac{dU}{d\lambda} U^{-1} \). We leave the details of the computation to the reader.
approximation of $V(T)$ provided that

$$v(t) = V(t), \quad (1.25)$$

$$\theta = \text{Tr} [\Sigma_\infty] > 0, \quad (1.26)$$

$$\kappa = -\frac{1}{\tau} \log \left( \frac{\mathbb{E} [V(T) | \Sigma(t)] - \theta}{V(t) - \theta} \right), \quad (1.27)$$

$$\eta^2 = \frac{\kappa \text{Var} [V(T) | \Sigma(t)]}{a \left( (1 - a) V(t) + \frac{a}{2} \theta \right)}, \quad (1.28)$$

with $a = (1 - \exp(-\kappa \tau))$ and provided that $V(t) \neq \theta$. A sufficient condition for $\kappa$ to be well-defined (i.e. to be a positive real number) is to have the matrix $F = M \Sigma(t) + \Sigma(t) M^\top + \beta Q^\top Q$ (positive or negative) definite.

**Proof.** Given the distributional properties of CIR process, the process $\tilde{v}(t)$ with parameters (1.25)-(1.28) has the same first two $T$-conditional moments of $V(T)$. Furthermore, given the stability\(^5\) of $M$, we have that $\theta \geq 0$, with the equality that holds only in the degenerate case $Q = 0_d$ (the zero matrix of order $d$). Indeed, for $M$ stable, the solution of the Lyapunov equation (1.2) admits the integral representation

$$\Sigma_\infty = \beta \int_0^\infty \exp (uM) Q^\top Q \exp (uM^\top) \, du$$

which is positive semidefinite and then its trace is non-negative. Additionally it is possible to show that $\text{Tr} [\Sigma_\infty] = 0$ if and only if $Q = 0_d$. To see this, let us write

$$\theta = \beta \text{Tr} \left[ \int_0^\infty \exp (uM) Q^\top Q \exp (uM^\top) \, du \right]$$

$$= \beta \text{Tr} \left[ Q^\top Q \int_0^\infty \exp (uM) \exp (uM^\top) \, du \right].$$

We now have to prove that $\theta$ is strictly positive. The claim follows\(^6\) by using the fact that $\theta$ is the trace of the product of positive semi-definite matrix $\beta Q^\top Q$ and the positive definite matrix $\int_0^\infty \exp (uM) \exp (uM^\top) \, du$.

From (1.27) we see that $\kappa \in \mathbb{R}_{>0}$ as soon as $0 < \frac{\mathbb{E} [V(T) | \Sigma(t)] - \theta}{V(t) - \theta} < 1$, with 2 possible cases:

5A square matrix is called stable if all its eigenvalues have strictly negative real part.

6Specifically, there is no nilpotent matrix (i.e. with all the eigenvalues equal to zero), except the zero matrix, that can be obtained as the product of 2 non-negative matrices. Moreover we exploit the fact that if $A$ and $B$ are two non-zero square matrices such that $AB = 0$, then both $A$ and $B$ must be singular. Given that $\int_0^\infty \exp (uM) \exp (uM^\top) \, du$ is positive definite (indeed, it is the integral of a positive definite matrix), we get the result.
1.2. Definition of Wishart process and basic properties

- if \( V(t) - \theta > 0 \), it must hold \( \theta < \text{Tr} [\Gamma(\tau) + \beta \Theta(\tau)] < V(t) \),
- otherwise \( \theta > \text{Tr} [\Gamma(\tau) + \beta \Theta(\tau)] > V(t) \geq 0 \).

In order for the condition to be fulfilled for any \( T > t \), we can require the function \( g : \tau \mapsto \text{Tr} [\Gamma(\tau) + \beta \Theta(\tau)] \) to be monotone in \([0, \infty)\). We indeed have that \( V(t) = \text{Tr}[\Gamma(0) + \beta \Theta(0)] \) and \( \theta = \lim_{\tau \to +\infty} \text{Tr}[\Gamma(\tau) + \beta \Theta(\tau)] \). It is easy to check that \( E(\tau) = \mathbb{E}[\Sigma(T)|\Sigma(t)] = \Gamma(\tau) + \beta \Theta(\tau) \) is solution of the first order, linear inhomogeneous matrix ODE

\[
\dot{E}(\tau) = ME(\tau) + E(\tau)M^T + \beta Q^T Q
\]

with \( E(0) = \Sigma(t) \). Differentiating (1.29) we obtain

\[
\dot{E}(\tau) = M \dot{E}(\tau) + \dot{E}(\tau)M^T
\]

with initial condition \( \dot{E}(0) = M \Sigma(t) + \Sigma(t)M^T + \beta Q^T Q = F \). Following standard variation of constants arguments, we have

\[
\dot{E}(\tau) = \exp(\tau M)F \exp(\tau M^T).
\]

This in turns means that the derivative of function \( g \) can be written as

\[
\frac{d}{d\tau} g(\tau) = \text{Tr} \left[ \dot{E}(\tau) \right] = \text{Tr} \left[ F \exp(\tau M^T) \exp(\tau M) \right].
\]

As shown in [92], for any real symmetric matrix \( A \) and \( B \in \mathcal{S}_d^+ (\mathbb{R}) \), the following inequality holds

\[
a_i \text{Tr}[B] \leq \text{Tr}[AB] \leq a_i \text{Tr}[B]
\]

where \( a_i \) is the \( i \)-th largest eigenvalue of \( A \). Applying this result to (1.31) we get the sufficient condition for the monotonicity of \( g(\tau) \) in terms of definiteness of \( F \). We see, indeed, that if \( F \) is (positive or negative) definite, the derivative of \( g \) is bounded to be always positive or negative - depending on the sign of the eigenvalues of \( F \) - on the interval \([0, \infty)\).

Unfortunately, since we deal with traces of matrix products, obtaining sharp conditions for the monotonicity of \( g \) is not an easy task. We are able just to provide the sufficient condition claimed, based on the known (loose) bounds for the trace of the product of two matrices. However, for a wide range of realistic parameter values we could have that \( \kappa \) is still well-defined even if the condition on \( F \) is not fulfilled. As an example, we consider the parameters set:

\[
\beta = 1.1, \quad \Sigma_0 = \begin{bmatrix} 0.0298 & 0.0119 \\ 0.0119 & 0.0108 \end{bmatrix}, \quad (1.32)
\]
Chapter 1. Wishart Processes in Finance

Figure 1.1: Left panel: $g(\tau)$ in the interval $[0, 4]$ (black line). It departs from $\text{Tr} [\Sigma_0]$ (blue line) for $\tau = 0$ and converges towards $\theta$ (red line) as $\tau \to +\infty$. Right panel: derivative of $g(\tau)$ (black line) and its bounds.

$$M = \begin{bmatrix} -1.2479 & -0.8985 \\ -0.0820 & -1.1433 \end{bmatrix}, \quad Q = \begin{bmatrix} 0.3417 & 0.3493 \\ 0.1848 & 0.3090 \end{bmatrix}. \quad (1.33)$$

The parameters are taken from [32] with the exception of $\beta$ that we set higher than 1 in order to deal with a well-defined Wishart process. Interestingly, this is the same set used in [70]. As shown in Figure 1.1, the function $g(\tau)$ is monotonic even if $\mathcal{F}$ is indefinite (its eigenvalues are 0.3106 and $-0.0278$). From the rightmost panel of Figure 1.1, it is also evident that for short time frames, the bounds for the trace of $\dot{E}(\tau)$ are not really tight. For $d = 2$, this means that if the non-dominant eigenvalue has different sign with respect to the dominant one but its value is enough close to zero, the sign of the derivative of $g(\tau)$ is likely not to change.

We now take a look at the distribution of the square root of $V(T)$ that can be used, as shown in the next section, to model the stochastic volatility in a multi-factor extension of the Heston model. It is possible to recover the probability density of $\sqrt{V(T)}$ by applying the formula

$$p_{\sqrt{V(T)}}(x) = 2x p_{V(T)} \left( x^2 \right) \quad \forall x \in \mathbb{R}_{\geq 0},$$

where the probability density function of $V(T)$, $p_{V(T)}$, can be obtained via FFT or quadrature methods from (1.21). In Figure 1.2 we also consider the distribution induced by the two approximations taken into account. It is worth noting that for the affine transformation approach, there is a discontinuity at the point $\sqrt{\alpha_0}$ such that for lower values the probability density is zero: the far left tail of the exact distribution is
1.2. Definition of Wishart process and basic properties

Figure 1.2: Comparison among exact probability density function of $\sqrt{V(T)}$ (black line), approximation via affine transformation of non-central chi-square random variables (1.24) (red line) and our moment-matching approximation (blue line) for different time horizons.

unattainable and compensated by a corresponding peak in a (positive) neighborhood of the discontinuity. Our moment-matching approximation, on the other hand, does not suffer from this drawback by construction (since it has no displacement term) and it is still capable of reproducing the distributional properties of $\sqrt{V(T)}$. In particular, our approach gives an accurate approximation in the right tail of the distribution (comparable to the alternative affine approximation) while for intermediate values it fails in some extent in reproducing correctly the peak in the exact distribution. Considering,
Chapter 1. Wishart Processes in Finance

however, that this is a approximation based simply on the matching of the first 2 moments of $V(T)$, a lack of accuracy is somehow expected. Finally, we want to stress that the approximation proposed in Proposition 4 is not meant to strictly describe the distribution of $V(T)$ but rather to provide a simple and reliable tool to map the parameters that drive the trace of $\Sigma(t)$ to those of a simpler scalar CIR process. In the next section we use such mapping to develop efficient numerical techniques for models that use $\sqrt{V(t)}$ to describe the multi-factor dynamics of asset volatility.

In the following sections we analyze in detail the existing models that make use of Wishart process as multidimensional driving factor for asset(s) volatility. In particular, we focus our attention to the Wishart Stochastic Volatility model in [34] and the Wishart Affine Stochastic Correlation model in [33], with the latter that represents a generalization of the multi-asset model presented in [55]. Since our main concern in this paper is the pricing of contingent claims, from now on we assume to operate under a risk-neutral measure as defined in the standard way.

1.3 The Wishart Stochastic Volatility Model

In order to overcome inherent limitations of 1-factor SV models in describing the term structure of volatility skew, as documented for example in [26] and [24], a matrix generalization of Heston model is proposed in [34]. This is the Wishart Stochastic Volatility model (WSVM) where the dynamics of the forward-price of an equity asset is given by

$$df(t) = f(t) \operatorname{Tr}\left[ \sqrt{\Sigma(t)} dB(t) \right], \quad f(0) > 0,$$

(1.34)

where $B(t)$ is a $d$-dimensional matrix Brownian such that

$$B(t) = W(t) R^T + Z(t) \sqrt{I_d - RR^T},$$

(1.35)

with $Z(t)$ another matrix of Brownian motions independent of $W(t)$ and $R \in \mathcal{M}_d(\mathbb{R})$ that fulfills the condition $I_d - RR^T \in \mathcal{S}_d^+(\mathbb{R})$. This correlation structure is required in order to preserve the analytical tractability of the model: the WSVM so defined, indeed, belongs to the class of affine models in the sense of [41].

From now on we consider the WSVM dynamics (1.34) with the underlying assumption of $d = 2$. This is the most common parametrization appeared in literature and constitutes an adequate balance between parsimony and flexibility.

In the original paper, the authors show that the WSVM can be expressed in a scalar form that, as we are going to see, turns out to be extremely useful in our framework. Before stating this result, we recall an auxiliary Lemma:

Lemma 1 (Lemma 4.6 in [81]). Let $X(t)$ be a continuous stochastic process on $\mathcal{S}_d^+(\mathbb{R})$ and let $h : \mathcal{M}_d(\mathbb{R}) \to \mathcal{M}_d(\mathbb{R})$. Then there exists a scalar Brownian motion $w_h(t)$ such
that

\[
\text{Tr} \left[ \int_0^T h(X(u))dW(u) \right] = \int_0^T \sqrt{\text{Tr} [h(X(u))^	op h(X(u))]}dw_h(u)
\] (1.36)

holds true.

**Proposition 5 (Scalar version of WSVM dynamics).** Let \( y(t) = \log(f(t)) \) be the asset log-price, then its dynamics in the WSVM can be written as

\[
dy(t) = -\frac{1}{2} \mathcal{V}(t)dt + \sqrt{\mathcal{V}(t)} \, dB(t),
\] (1.37)

\[
d\mathcal{V}(t) = \left( \text{Tr} \left[ \beta Q^	op Q \right] + 2 \text{Tr} \left[ M \Sigma(t) \right] \right) dt + 2 \sqrt{\text{Tr} [\Sigma(t)Q^	op Q]} \, dw(t),
\] (1.38)

where \( b(t) \) and \( w(t) \) are two scalar Brownian motions with stochastic correlation given by

\[
\rho_W(t) = \frac{\text{Tr} [RQ\Sigma(t)]}{\sqrt{\text{Tr} [\Sigma(t)]} \sqrt{\text{Tr} [Q^	op Q\Sigma(t)]}}.
\] (1.39)

**Proof.** From standard Itô’s lemma we immediately have that

\[
dy(t) = -\frac{1}{2} \mathcal{V}(t)dt + \text{Tr} \left[ \sqrt{\Sigma(t)} \, dB(t) \right]
\] (1.40)

and applying Lemma 1 we get

\[
\text{Tr} \left[ \sqrt{\Sigma(t)} \, dB(t) \right] = \sqrt{\text{Tr} [\Sigma(t)]} \, db(t) = \sqrt{\mathcal{V}(t)} \, db(t).
\] (1.41)

An analogous argument applies for the dynamics of \( \mathcal{V}(t) \) whose stochastic term in (1.19) reads

\[
\text{Tr} \left[ \sqrt{\Sigma(t)} dW(t)Q \right] = \sqrt{\text{Tr} [\Sigma(t)Q^	op Q]} \, dw(t).
\] (1.42)

The expression for \( \rho_W(t) \) directly arises from the covariation of \( b(t) \) and \( w(t) \):

\[
d[b(t), w(t)] = \frac{\text{Tr} \left[ \sqrt{\Sigma(t)} dW(t) R \right]}{\sqrt{\text{Tr} [\Sigma(t)]}} \frac{\text{Tr} \left[ \sqrt{\Sigma(t)} dW(t)Q \right]}{\sqrt{\text{Tr} [\Sigma(t)Q^	op Q]}}
\]

\[
= \frac{\text{Tr} [RQ\Sigma(t)]}{\sqrt{\text{Tr} [\Sigma(t)]} \sqrt{\text{Tr} [Q^	op Q\Sigma(t)]}} dt
\]

where we refer to [34] for the complete computations.

This result highlights also the appealing property that, in the WSVM framework, the correlation between stock returns and volatility is stochastic.

An important peculiarity of WSVM is that, differently than other multifactor extensions of the Heston model, like the Bi-Heston model in [24], matrix specification of volatility...
Chapter 1. Wishart Processes in Finance

factors allows to separately manage and calibrate implied volatility levels and skew thanks to the presence of non null off-diagonal elements in \( \Sigma_0 \). Let us suppose to fix the elements on the diagonal of \( \Sigma_0 \) to match the term structure of implied volatility. We still have the possibility to fit the skew thanks to the residual elements in \( \Sigma_0 \). This is also confirmed by the analysis performed in [32] where the expansion of model implied volatility for short times to maturity is found to be

\[
\sigma_{imp}^2 = \text{Tr} [\Sigma_0] + \frac{\text{Tr} [RQ \Sigma_0]}{\text{Tr} [\Sigma_0]} m f - \frac{m f^2}{(\text{Tr} [\Sigma_0])^2} \left( \frac{1}{3} \text{Tr} [Q^\top Q \Sigma_0] + \frac{1}{3} \text{Tr} [RQ (Q^\top R^\top + RQ) \Sigma_0] - \frac{5}{4} \frac{(\text{Tr} [RQ \Sigma_0])^2}{\text{Tr} [\Sigma_0]} \right)
\]

(1.43)

where \( m f = \log \frac{K_f(t)}{f(t)} \) denotes the log-forward moneyness. From (1.43), we can appreciate that the off-diagonal element \( \Sigma_{12} \) does not affect the level of the smile but it has a relevant impact on the slope of implied volatility.

1.3.1 Connection with Heston and Bi-Heston model

In this section we aim at studying the relationship between WSVM and other (simpler) affine stochastic volatility models, namely the Heston [63] and Bi-Heston [24] ones. The idea comes from the distributional properties of the trace of Wishart process and its role in describing the asset volatility. This will turn out to be extremely useful in devising efficient calibration algorithms. In Section 1.2.2 we propose to approximate the conditional distribution of \( V(T) \) by means of a scaled non-central \( \chi^2 \) random variable and we derive the (\( T \)-specific) parameters of the corresponding CIR process. For a fixed time horizon \( T \), then, the WSVM dynamics can be approximated by the Heston one

\[
\begin{align*}
 df(t) &= f(t) \sqrt{v(t)} \, db(t) \\
 dv(t) &= \kappa (\theta - v(t)) dt + \eta \sqrt{v(t)} dw(t), \quad v(0) = v_0 \geq 0
\end{align*}
\]

with parameters \( v_0, \kappa, \theta \) and \( \eta \) defined in Proposition 4 (here we assume \( t = 0 \)). The asset-volatility correlation is driven by a constant parameter \( \rho \) (in the sense that \( db(t) dw(t) = \rho dt \)) that we set as

\[
\rho = \frac{\text{Tr} [RQ \Sigma_0]}{\sqrt{\text{Tr} [\Sigma_0] \text{Tr} [Q^\top Q \Sigma_0]}},
\]

(1.44)

where the right-hand side of (1.44) is the initial value of the process \( \rho_W(t) \). This means that given a WSVM parameters set \( \pi_W \) and a time horizon \( T \), we can construct a model approximation by mapping WSVM parameters into Heston ones. In other words, we define a function \( g^{H-W} : \mathbb{R}^{N_W} \times \mathbb{R}_{>0} \to \mathbb{R}^5 \) with \( N_W \) the number of parameters in the
1.3. The Wishart Stochastic Volatility Model

chosen configuration of WSVM such that \( g_{H-W}(\pi_W, T) = \pi_H = [v_0, \kappa, \theta, \eta, \rho]^\top \). We refer to function \( g_{H-W} \) as the WSVM-Heston mapping.

From Corollary 3 we know that the asset variance in WSVM is described by a linear combination of 2 independent non-central \( \chi^2 \) random variables. This is the same distributional assumption underlying the Bi-Heston model proposed in [24], where the asset dynamics is the following:

\[
\begin{align*}
    df(t) &= f(t) \left( \sqrt{v_1(t)} dB_1(t) + \sqrt{v_2(t)} dB_2(t) \right) \\
    dv_1(t) &= \kappa_1 (\theta_1 - v_1(t)) dt + \eta_1 \sqrt{v_1(t)} dW_1(t), \quad v_1(0) = v_{0,1} \geq 0 \\
    dv_2(t) &= \kappa_2 (\theta_2 - v_2(t)) dt + \eta_2 \sqrt{v_2(t)} dW_2(t), \quad v_2(0) = v_{0,2} \geq 0
\end{align*}
\]

with \( dB_i(t)dW_i(t) = \rho_i dt \) for \( i = 1, 2 \) and all other correlation are equal to zero to preserve the affinity of the model. As shown in [24], in this model the variance of log-asset price is the sum of the 2 independent CIR processes

\[
\text{Var} [d \log(f(t))] = (v_1(t) + v_2(t)) dt = v_{BH}(t) dt. \quad (1.45)
\]

Furthermore, the model presents a stochastic asset-variance correlation given by

\[
\text{Corr} [d \log(f(t)), dv_{BH}(t)] = \frac{\eta_1 \rho_1 v_1(t) + \eta_2 \rho_2 v_2(t)}{\sqrt{\eta_1^2 v_1(t) + \eta_2^2 v_2(t)}} dt. \quad (1.46)
\]

In the lights of the analogy between (1.20) and (1.45), it could be interesting to find a suitable parameters set for the Bi-Heston model such that it represents a close approximation of WSVM.

For a fixed \( T > 0 \), we propose to approximate the WSVM dynamics (1.34) by means of a Bi-Heston model whose CIR processes parameters are (for \( i = 1, 2 \))

\[
\begin{align*}
    v_{0,i} &= \tilde{v}_i, \quad (1.47) \\
    \kappa_i &= -\frac{1}{T} \log \left( \frac{\zeta_i}{\tilde{v}_{0,i}} \right), \quad (1.48) \\
    \eta_i &= 2 \sqrt{\frac{\varepsilon_i \kappa_i}{(1 - e^{-\kappa_i T})}}, \quad (1.49) \\
    \theta_i &= \frac{\beta_i \eta_i^2}{4 \kappa_i} \quad (1.50)
\end{align*}
\]

where \( \tilde{v}_i \) and \( \zeta_i \) are, respectively, the \( i \)-th diagonal elements of matrices \( Q^\top \Sigma_0 Q \) and \( Q^\top \Gamma(T) Q \).

For example, by considering \( M, Q \) and \( R \) to be full matrices we have \( \pi_W = [\beta, \Sigma_{11}, \Sigma_{12}, \Sigma_{22}, M_{11}, M_{12}, M_{21}, M_{22}, Q_{11}, Q_{12}, Q_{21}, Q_{22}, R_{11}, R_{12}, R_{21}, R_{22}]^\top \) and \( N_W = 16 \).
Parameters $\kappa_i$, $\theta_i$, and $\eta_i$ directly follow from the representation of the distribution of $V(T)$ as formulated in Corollary 3. Reasonably, we set the initial values of variance processes (1.47) to be equal to the diagonal elements of the matrix $Q^\top \Sigma_0 Q$, i.e., the matrix obtained applying the change of basis that diagonalizes $\Theta(T)$.

The choice of coefficients $\rho_i$ is the most problematic. Not only because we want to "map" the effect of the $2 \times 2$ matrix $R$ into just 2 parameters, but also because there seems not to be an immediate way to link (1.39) and (1.46). In [34] it is shown that the covariation between $\Sigma_i$ and log-asset price $y$ induced by the WSVM is $d[y(t), \Sigma_i(t)] = 2 \sum_{k,h=1}^2 \Sigma_{ih}(t)Q_{ki}R_{hk}dt = 2Dy\Sigma_i(t)dt$. A tempting solution could be, then, to set $\rho_i$ as the stochastic correlation between $\Sigma_i$ and $y$ valued at $t = 0$:

$$\rho_i = \frac{Dy\Sigma_i(0)}{\sqrt{\text{Tr}[\Sigma_0]\sqrt{\Sigma_i}Q_{ii}^*}}$$

where we make use of (1.4) to obtain $d[\Sigma_i(t), \Sigma_i(t)] = 4\Sigma_i(t)Q_{ii}^*dt$. However, extensive numerical experiments provide evidences in favour of the following alternative formulation

$$\rho_i = \frac{Dy\Sigma_i(0)}{\Sigma_i\sqrt{Q_{ii}^*}}.$$

Unfortunately, we do not have any formal justification for this formula other than the heuristics. Notwithstanding, we suppose that a more reliable parametrization of $\rho_i$ is possible and we leave the impact of this choice to further research. As done for the Heston model, we can define the WSVM-BiHeston mapping $g^{BH-W}(\pi_W, T) = \pi_{BH} = [v_0, 1, \kappa_1, \theta_1, \eta_1, \rho_1, v_{0,2}, \kappa_2, \theta_2, \eta_2, \rho_2]^T$.

We now want to test the approximations proposed and assess their accuracy. In the first numerical experiment, we compare the performance of our new methodologies with the second-order price expansion developed in [49]. The parameters are those considered in [49]:

$$\beta = 4, \quad \Sigma_0 = \begin{bmatrix} 0.4 & 0 \\ 0 & 0.35 \end{bmatrix}, \quad M = \begin{bmatrix} -1 & 0 \\ 0 & -0.8 \end{bmatrix}, \quad Q = \begin{bmatrix} 0.3 & 0 \\ 0 & 0.2 \end{bmatrix}, \quad R = aI_2$$

for 2 different choices of $a$: 0 and $-0.5$. The constraint of $R$ to be a multiple of the identity matrix is, indeed, a strict requirement for the derivation of the formula in [49]. The results corresponding to $T = 1$ and $r = 3\%$ are shown in Figure 1.3. In the simple case of zero asset-volatility correlation (leftmost panel), all the methodologies considered work extremely well for any level of moneyness. Introducing a negative correlation, though, we notice a deterioration in the accuracy of the second-order expansion for deep in-the-money options (rightmost panel). On the other hand, the Heston and Bi-Heston approximations are still hardly distinguishable from the true values. Another
1.3. The Wishart Stochastic Volatility Model

Figure 1.3: Implied variance smiles for the WSVM (Exact), the second-order expansion in [49] (GP), our Bi-Heston (BH) and Heston (H) approximations. Left panel $a = 0$, right panel $a = -0.5$.

advantage of the new approaches proposed is their extreme simplicity: we just need to price options in well-known affine models, while the formula in [49] requires to compute several integrals of (products of) exponential and hyperbolic functions of matrix argument. Additionally, if used to tackle the calibration problem, our techniques lead to the explicit computation of the gradient of the objective function. This topic will be further extended in the next Section. Moving towards more challenging parametrizations, we investigate the dataset (1.33) coupled with matrix

$$ R = \begin{bmatrix} -0.2243 & -0.1244 \\ -0.2545 & -0.7230 \end{bmatrix} $$

as originally calibrated in [32]. In this case, the approximation in [49] is no longer applicable since $R \neq aI_2$. On the contrary, our methodologies do not suffer from this limitation and we report the results in Figure 1.4 for different time horizons. As we consider longer maturities, not surprisingly, we experience a worsening in the accuracy of both approximations. The (full) matrix structure of WSVM parameters is too complex to be entirely captured. Notwithstanding, the smile generated by the Bi-Heston approach is still reasonably in line with the WSVM values: in the worst-performing case ($T = 2$ years) the mean absolute error in volatility terms over the range of moneyness $[40\% - 140\%]$ is $0.9\%$.

1.3.2 Characteristic function and calibration to market prices

Given the affinity of WSVM, it is possible to express the characteristic function of log-prices $y(T)$ as the exponential of an affine combination of state variables $y(t)$ and $\Sigma(t)$. In particular the closed formula for the characteristic function is derived in [34] where
the resulting matrix Riccati equation is solved via linearization technique. Without entering into technical details (for which we refer the interested reader to the original paper [34]) we focus our attention on the formula as presented in the following Proposition.

**Proposition 6 (Characteristic function of log-price in WSVM).** Let the log-forward price \( y(t) \) be described by (1.40) and \( \lambda \) be an auxiliary real variable (different from those used above). Then for \( T > t \), the WSVM (conditional) characteristic function of \( y(T) \) admits the following closed formula representation

\[
\phi^W_y(\lambda, \tau) = \mathbb{E} \left[ \exp (\imath \lambda y(T)) \mid y(t) \right] = \exp (\imath \lambda y(t) + \text{Tr} \left[ A_y(\tau) \Sigma(t) \right] + b_y(\tau)), \tag{1.52}
\]
1.3. The Wishart Stochastic Volatility Model

with the deterministic matrix $A_y(\tau)$ and the scalar function $b_y(\tau)$ given by

$$A_y(\tau) = A_{22}(\tau)^{-1} A_{21}(\tau),$$

$$b_y(\tau) = -\frac{\beta}{2} \text{Tr} \left[ \log(A_{22}(\tau)) + \tau (M + i\lambda Q^\top R^\top) \right],$$

and

$$
\begin{bmatrix}
A_{11}(\tau) & A_{12}(\tau) \\
A_{21}(\tau) & A_{22}(\tau)
\end{bmatrix}
= \exp \left( \tau \begin{bmatrix}
M + i\lambda Q^\top R^\top & -2Q^\top Q \\
\frac{i\lambda(\lambda-1)}{2} I_d & (M + i\lambda Q^\top R^\top)^\top
\end{bmatrix} \right).
$$

Proof. See [34].

Given the availability of a closed formula for the characteristic function of $y(T)$, we can price plain vanilla options through efficient numerical techniques. However, from a computational point of view, a direct application of (1.40) for calibration purposes could turn out to be highly cumbersome. Firstly because we are required to compute several functions of matrix argument each time, and secondly because the corresponding optimization problem would present a multiplicity of local minima and a strong dependence on the starting point. It is also important to point out that, as stated in [32], the use of existing approximations (both in terms of price and implied volatility) is likely to be restricted to narrow ranges around the at-the-money level. In the light of the above, then, efficient calibration of WSVM is still an open problem that could severely limit the real-world application of the model.

To overcome such limitations, we propose a fast and accurate calibration procedure that relies on the model approximations developed in Section 1.3.1. We start by illustrating the properties of the optimization problem. Let $C(K,T)$ be the market price of a call option struck at $K$ with maturity in $T$ years and $C_{\text{Model}}(\pi, K, T)$ be the corresponding price obtained via the chosen model with (model-specific) parameters set $\pi$. We formulate the calibration problem as an inverse problem of the form

$$
\min_{\pi \in \mathbb{R}^{N_\pi}} \frac{1}{2} f_{\text{obj}}(\pi) \quad (1.53)
$$

where $N_\pi$ is the dimension of the parameters set and $f_{\text{obj}}$ is the so-called objective function, i.e. the metric that defines the distance between market and model values. Several specifications for $f_{\text{obj}}$ can be used and it is well understood that they can lead to quite different results. In our approach we follow [32] and define

$$
f_{\text{obj}}(\pi) = \sum_{s=1}^{N_s} \left( \frac{C_{\text{Model}}(\pi, K_s, T_s) - C(K_s, T_s)}{\omega_s} \right)^2, \quad (1.54)
$$

that is, we consider the sum of the square weighted difference between model and market values over a set of $N_s$ quoted instruments. The weights $\omega_s$ can be chosen.
in order to put more emphasis on a certain subset of options (in terms, for example, of liquidity, moneyness or time to maturity). In the following we consider $\omega_s$ to be the inverse of (squared) Black-Scholes vega computed with respect to the $s$-th market option price. This choice of weights put more emphasis on short dated OTM options that otherwise would have almost no influence in the calibration procedure.

Defining the (weighted) residuals $\tilde{r}_s(\pi) = \left( C_{Model}(\pi, K_s, T_s) - C(K_s, T_s) \right) / \omega_s$, the calibration problem can be written as

$$\min_{\pi \in \mathbb{R}^{N_\pi}} \frac{1}{2} f_{obj}(\pi) = \min_{\pi \in \mathbb{R}^{N_\pi}} \frac{1}{2} \tilde{r}(\pi)^\top \tilde{r}(\pi)$$

with $\tilde{r}(\pi) \in \mathbb{R}^{N_s}$. Furthermore, let $J_{Model}(\pi) = \nabla \tilde{r}(\pi)^\top \in \mathbb{R}^{N_s \times N_\pi}$ be the Jacobian matrix of the residuals vector $\tilde{r}(\pi)$ with elements

$$J_{r,s} = \frac{\partial \tilde{r}_s(\pi)}{\partial \pi_r} = \frac{1}{\omega_s} \frac{\partial C_{Model}(\pi, K_s, T_s)}{\partial \pi_r},$$

then the gradient of the objective function is given by $\nabla f_{obj} = J_{Model}(\pi) \tilde{r}(\pi)$. Usually, for many financial models the gradient of $f_{obj}$ is computed numerically by, for example, finite differences thus requiring a large number of function evaluations. There is, however, a noticeable exception: in [30], the authors exploit an alternative representation of the Heston model characteristic function to obtain the analytical gradient of the price of a vanilla option with respect to model parameters (with our notation, this means that $J^H(\pi_W)$ can be computed explicitly, where $H$ stands for Heston model). The resulting calibration algorithm is extremely fast and robust.

We propose to combine the algorithm in [30] with the model approximations developed in section 1.3.1. In other words, we transform, by parameters mapping, the WSVM calibration problem in a simpler one for which we are able to compute the gradient of the objective function in closed formula. The resulting procedure has two inherent advantages from a computational point of view: firstly we replace the WSVM characteristic function with a less computational demanding one; secondly we reduce the number of function evaluations (we do not need to approximate the gradient via finite differences). More importantly, by letting the exact gradient to drive the optimization routine, we are able to rapidly identify a suitable parameters set in the (high dimensional) search space.

We show how to compute the gradient of the objective function. Let us consider firstly the WSVM-Heston case: the (transformed or approximated) calibration problem can be written as

$$\min_{\pi_W \in \mathbb{R}^{N_W}} \frac{1}{2} f_{obj}(\pi_W) = \min_{\pi_W \in \mathbb{R}^{N_W}} \frac{1}{2} \tilde{r}(g^H-W(\pi_W))^\top \tilde{r}(g^H-W(\pi_W))$$

where, for simplicity, we suppress the time dependency. By chain rule, we have that the Jacobian matrix of residuals with respect to WSVM parameters is given by

$$J^W(\pi_W) = J^H-W(\pi_W) J^H(g^H-W(\pi_W))$$

8We remark, however, that the computation of function $h$ must be performed for any maturity in the calibration basket.
where \( J^H \) is known thanks to [30]. The matrix \( J^{H-W}(\pi_W) = \nabla g^{H-W}(\pi_W) \in \mathbb{R}^{N_W \times 5} \) is the Jacobian matrix of function \( g^{H-W} \) with respect to WSVM parameters, with elements
\[
J_{q,r}^{H-W} = \frac{\partial g_{r}^{H-W}(\pi_W)}{\partial \pi_{W,q}}.
\]
(1.59)
that, as shown in Appendix 1.A.2 can be computed in closed formula. Analogously to the general case, finally, it holds that \( \nabla f_{obj} = J^{W}(\pi_W) \tilde{r}(g^{H-W}(\pi_W)) \).

Straightforwardly, we can apply the same methodology in conjunction with the Bi-Heston approximation. In Appendix 1.A.3 we show how to extend the approach in [30] in order to compute the gradient of Bi-Heston call options with respect to model parameters. We then couple the matrix \( J^{BH} \) with the map \( g^{BH-W} \) to write, as in the previous case:
\[
J^{W}(\pi_W) = J^{BH-W}(\pi_W)J^{BH}(g^{BH-W}(\pi_W, T)),
\]
(1.60)
where, once more, \( J^{BH-W} \) can be obtained in closed formula (see Appendix 1.A.4).

In the light of the evidences shown in the previous section, we decide to implement the Bi-Heston approximation when it comes to calibrate WSVM to market data. Nonetheless, we are confident that both methodologies can be helpful in understanding the role of WSVM parameters and performing models comparisons in the spirit of [32]. We devise a 2 steps calibration procedure: initially we make use of WSVM-BiHeston mapping to tackle the simplified calibration problem and obtain a robust guess of the optimal parameters. In the second step, we consider this parameters set as the starting point of the standard calibration algorithm based on the pricing of vanilla options via WSVM characteristic function (1.52). This can be interpreted as the “fine-tuning” phase in order to further improve the accuracy achieved. By starting from a robust initial guess, we expect the algorithm to converge after few function evaluations. In this step, we implement the so-called COS method [45]. This approach basically exploits the relation between the characteristic function and the coefficients of the Fourier cosine series expansion of the corresponding probability density function.

We test the proposed algorithm on the calibration of WSVM parameters to market prices of 182 European call options written on DAX index as of February 3, 2016. We here consider matrix \( M \) to be symmetric: this choice is motivated from the fact that the resulting WSVM model embeds an higher degree of tractability. As shown in [49], for example, in this setting it is possible to compute explicitly the expected value of integrated variance, thus opening the way to consistent pricing of volatility derivatives. Calibrated parameters are shown in the leftmost column of Table 1.4 in Appendix 1.C.

---

9As far as we know, this is the first time that such a result is derived. It provides an efficient tool to calibrate Bi-Heston parameters to market data.
As we can see, the Bi-Heston step takes less than 2 seconds. It is worthwhile noting that the resulting parameters matrices are all diagonal. This is, however, not the standard case. In the second step, the fine-tuning effect is quite relevant: the diagonal elements are very close to those found with the Bi-Heston mapping while the off-diagonal ones are set so that the accuracy is significantly improved (both mean squared errors in price and volatility terms are roughly halved). The efficiency of the overall procedure can be also inferred from the fact that the second step takes only 24.33 seconds. In comparison, the calibration procedure fully based on the WSVM characteristic function would take roughly 250 seconds. Figure 1.5 shows results of calibration procedure in terms of corresponding implied volatility surface and absolute error with respect to market values. We see that apart from very short dated far from the money options the WSVM can fit the market surface quite well. According to the evidences in [32] the resulting value of parameter $\beta$ is strictly lower than 1.

In light of what we have shown in Proposition 1, the conditions for the existence and uniqueness of solution to SDE (1.1) are consequently not satisfied. This is a crucial restriction specifically when we use calibrated WSVM to price derivatives for which we need to rely on simulation methods (which is of course the main interest for a structuring team dealing with realistic applications). Even if we use an unbiased (or yet exact) method to simulate the Wishart process, we would not have any chance to consider the corresponding process as an approximation of the original SDE (1.1).

If our goal is to use WSVM in a Monte Carlo framework, then, we need to impose at least condition (1.7) that in the case $d = 2$ corresponds to $\beta \geq 1$. We then perform

---

The algorithms are implemented via Matlab code on a laptop PC with an Intel Core i7 CPU and 8 GB RAM.
1.3. The Wishart Stochastic Volatility Model

The Wishart Stochastic Volatility Model (WSVM) is a dynamic model for the instantaneous variance process. It is particularly useful in pricing financial derivatives where the variance is not constant over time. In this section, we discuss the calibration of WSVM over the same market set with the additional constraint on parameter $\beta$. As far as we know this is the first attempt to deal with such a constrained problem in a real market context. Once more, we implement the two steps procedure described above. The conclusions about the algorithm efficiency are the same: the combined approach significantly reduces the computational time required: the overall procedure takes now just 16 seconds. The mechanics is also very similar: the first step fastly produce a robust guess while the final step act as a targeted adjustment devoted to the enhancement of calibration accuracy. Also in this case, indeed, the calibrated parameters are very close to the outcomes of the first step. The output of the calibration routine is shown in Figure 1.6. By a direct comparison with Figure 1.5 we can easily see that the additional constraint on parameter $\beta$ has not negligible impact. In particular the accuracy for short dated far-from-the-money options is quite worsened.

The need for enhancing the pricing performance of WSVM for European claims while assuring that the variance process is well defined leads to the hybrid framework introduced in the next section.

1.3.3 New simulation schemes for the WSVM

The basic idea of our new schemes is to exploit the exact simulation method for the Wishart process developed in [3] to sample the WSVM log-price process (1.34). From now on, indeed, we assume to have a collection of $N$ simulated trajectories of the Wishart process over the time grid $0 = t_0 < t_1 < \ldots < t_{M_T} = T$ obtained via the exact scheme in [3]. For simplicity, we also consider a uniform time step $\Delta = T/M_T = t_{m+1} - t_m$ for $m = 0, 1, \ldots, M_T - 1$. The extension of such a scheme to the WSVM
Chapter 1. Wishart Processes in Finance

case is not however a trivial task since there is not an immediate way to reconstruct the correlation structure between the Wishart process and the asset price. We are not able then to substitute (1.35) into (1.34). This is a direct consequence of how the sampling procedure works: rather than simulating directly the desired Wishart process, the procedure is based on a sort of "bottom-up" approach. Basically the infinitesimal generator of a canonical Wishart process is found to be given by the sum of commuting operators associated to simple SDEs. Then any admissible Wishart process is linked to the canonical one through a law identity. This means that we need to simulate just the SDEs whose generators are linked to the one of the canonical Wishart process. As a consequence, we do not have a matrix of Gaussian variables $\hat{W}$ that can be used in (1.35) to discretize the trajectory of the asset price.

A method to circumvent this problem is proposed in [70] where the Wishart process is sampled exactly with the aforementioned technique and then the log-price is sampled from the conditional distribution $F_{y(T)\Sigma(T)}(y; s) = \Pr [y(T) \leq y|\Sigma(T) = s]$ that is retrieved numerically from the conditional characteristic function of $y$. However, as already pointed out in the Introduction, even if formally correct this approach turns out to be unfeasible in our case and in general when the construction of an entire path of $y$ is needed. It requires, indeed, the evaluation of special functions of matrix argument at each step of the sampling scheme.

We instead start from the intuition that the dynamics of $V(t)$ itself can be used to link (1.34) and (1.1) in a proper way. Let us combine (1.37) and (1.39) to rewrite the scalar dynamics of $y(t)$ as

$$dy(t) = -\frac{1}{2} V(t) dt + \sqrt{V(t)} (\rho W(t) dw(t) + \sqrt{1 - \rho^2 W(t)} dz(t))$$  \hspace{1cm} (1.61)

with $z(t)$ a scalar Brownian motion independent on $w(t)$. This will be the starting point of the two new simulation schemes we propose.

For the first approach, that we call the Stochastic Integral Approximation (SIA) scheme, we focus on the integral form of (1.61),

$$y(t + \Delta) = y(t) - \frac{1}{2} \int_t^{t+\Delta} V(s) ds + \int_t^{t+\Delta} \rho W(s) \sqrt{V(s)} dw(s) + \int_t^{t+\Delta} \sqrt{1 - \rho^2 W(s)} V(s) dz(s).$$  \hspace{1cm} (1.62)

In the discretization of (1.62), the time integral that involves $V$ can be approximated as in [4] by

$$\int_t^{t+\Delta} V(s) ds \approx \Delta (\mu_1 V(t) + \mu_2 V(t + \Delta))$$

where we can set, for example, $\mu_1 = \mu_2 = 0.5$. By Itô’s isometry we also have

$$\int_t^{t+\Delta} \sqrt{(1 - \rho^2 W(s)) V(s)} dz(s) \sim N \left( 0, \int_t^{t+\Delta} (1 - \rho^2 W(s)) V(s) ds \right)$$  \hspace{1cm} (1.63)
where the variance of the Gaussian random variable can be approximated as above. The only non-trivial term in (1.62) is the stochastic integral with respect to \( w \). Unfortunately, there is no possible way to formulate this integral in terms of the dynamics of \( V \). This is due to the fact that the stochastic component in (1.38) involves the trace of a matrix product of \( \Sigma(t) \) that cannot be decomposed as the product of the traces. However we can make use of the CIR approximation we devised in the previous section: if we write
\[
dV(t) \approx \kappa (\theta - V(t)) dt + \eta \sqrt{V(t)} dw(t)
\]
with parameters \( \kappa, \theta \) and \( \eta \) given in Proposition 4, we obtain
\[
\int_t^{t+\Delta} \sqrt{V(s)} dw(s) \approx \frac{1}{\eta} \left( V(t+\Delta) - V(t) - \kappa \theta \Delta + \kappa \int_t^{t+\Delta} V(s) ds \right).
\]
Let \( \hat{V}(t) \) and \( \hat{V}(t+\Delta) \) be the realizations of the trace of Wishart process for two adjacent time points as sampled with the scheme devised in [3], we propose the following discretization scheme for the log-price in WSVM:
\[
\hat{y}(t+\Delta) = \hat{y}(t) + K_0(t) + K_1(t) \hat{V}(t) + K_2(t) \hat{V}(t+\Delta)
\]
\[
+ \sqrt{K_3(t) \hat{V}(t) + K_4(t) \hat{V}(t+\Delta)} \hat{z}
\]
with \( \hat{z} \) a random number sampled from a \( N(0,1) \) and coefficients
\[
K_0(t) = -\frac{\kappa \theta \Delta}{\eta} \hat{\rho}_W(t),
\]
\[
K_1(t) = \mu_1 \Delta \left( \frac{\kappa}{\eta} \hat{\rho}_W(t) - \frac{1}{2} \right) - \frac{\hat{\rho}_W(t)}{\eta},
\]
\[
K_2(t) = \mu_2 \Delta \left( \frac{\kappa}{\eta} \hat{\rho}_W(t) - \frac{1}{2} \right) + \frac{\hat{\rho}_W(t)}{\eta},
\]
\[
K_3(t) = \mu_1 \Delta \left( 1 - \hat{\rho}_W^2(t) \right),
\]
\[
K_4(t) = \mu_2 \Delta \left( 1 - \hat{\rho}_W^2(t + \Delta) \right),
\]
where we perform the freezing of \( \rho_W(s) \) to write
\[
\int_t^{t+\Delta} \rho_W(s) \sqrt{V(s)} dw(s) \approx \rho_W(t) \int_t^{t+\Delta} \sqrt{V(s)} dw(s)
\]
that can be justified, as we will see, from the fact that the proposed scheme is mainly devoted to sampling the trajectories of \( y \) over fine time grids. The complete scheme is illustrated in Algorithm 1.

We also present a simpler simulation technique, indicated as Gaussian Variable Approximation (GVA) scheme, for the discretization of (1.61). Given a discrete trajectory
Chapter 1. Wishart Processes in Finance

Algorithm 1 The Stochastic Integral Approximation (SIA) scheme for the WSVM

1. Set $T > 0$, compute parameters $\kappa$, $\theta$ and $\eta$ as in Proposition 4 conditionally on $\Sigma$ with $\tau = T$
2. for each simulation trial $n, n = 1, \ldots, N$ do
3. Initialize $\hat{\Sigma}(0) = \Sigma_0$
4. Initialize $\hat{y}(0) = \log (f(0))$
5. for each time-step $t_m, m = 0, \ldots, M_T - 1$ do
6. Compute $\hat{\rho}_W(t_m)$ as given by (1.35)
7. Sample the Wishart process for a time step of $\Delta$ with initial state $\hat{\Sigma}(t_m)$ using the scheme in [3]
8. Compute $\hat{\rho}_W(t_m + \Delta)$ as given by (1.35)
9. Compute coefficients $K_0(t_m), \ldots, K_4(t_m)$ as in (1.68)-(1.72)
10. Draw $\hat{z} \sim N(0, 1)$
11. Discretize (1.62) by means of (1.67)
12. end for
13. end for

of $\mathcal{V}$, the Euler approximation of (1.61) reads

$$
\hat{y}(t + \Delta) = \hat{y}(t) - \frac{1}{2} \hat{V}(t) \Delta + \hat{\rho}_W(t) \sqrt{\hat{V}(t) \hat{w}} + \sqrt{(1 - \hat{\rho}_W^2(t)) \hat{V}(t) \tilde{z}} \quad (1.74)
$$

where $\hat{w}$ is the Gaussian random variable that would drive the discretization of (1.38) and $\tilde{z}$ is a Gaussian random variable with variance $\Delta$ and independent on $\hat{w}$. By exploiting the sampling technique in [3] to get $\hat{\Sigma}(t)$ and $\hat{\Sigma}(t + \Delta)$, we can use (1.38) to obtain an approximation of $\hat{w}$. Indeed, it holds that

$$
\hat{V}(t + \Delta) - \hat{V}(t) \approx \left( \text{Tr} \left[ \beta Q^\top Q \right] + 2 \text{Tr} \left[ M \hat{\Sigma}(t) \right] \right) \Delta(t) + 2 \sqrt{\text{Tr} \left[ \hat{\Sigma}(t) Q^\top Q \right]} \hat{w} \quad (1.75)
$$

from which we can easily retrieve the value of $\hat{w}$. Finally we can plug $\hat{w}$ into (1.74) to discretize the path of $y$. The complete procedure is summarized in Algorithm 2. The two schemes proposed share the remarkable property that, thanks to the scalar representation, we can avoid the generation of $d^2 - 1$ additional random variables for the dynamics of $y$ with an evident reduction of complexity burden.

We perform extensive numerical tests using the set of calibrated parameters obtained imposing $\beta \geq 1$ and reported in the rightmost column of Table 1.4 The exercise considered is the pricing of European call options with moneyness in the range $\{70\%, 100\%, 130\%\}$ and maturities varying from 6 months up to 3 years. All the numerical results are presented in Appendix 1.D.1. Despite the limited number of simulated trajectories, the

\[\text{As for example exotic options embedded in structured products with daily (or continuous) monitoring of underlying asset price.}\]
1.3. The Wishart Stochastic Volatility Model

Algorithm 2 The Gaussian Variable Approximation (GVA) scheme for the WSVM

1: for each simulation trial \( n, n = 1, \ldots, N \) do
2: \( \text{Initialize } \hat{\Sigma}(0) = \Sigma_0 \)
3: \( \text{Initialize } \hat{y}(0) = \log (f(0)) \)
4: for each time-step \( t_m, m = 0, \ldots, M_T - 1 \) do
5: \( \text{Compute } \hat{\rho}_W(t_m) \text{ as given by (1.35)} \)
6: \( \text{Sample the Wishart process for a time step of } \Delta \text{ with initial state } \hat{\Sigma}(t_m) \) using the scheme in [3]
7: \( \text{Approximate the Gaussian random variable } \tilde{w} \text{ by means of (1.75)} \)
8: \( \text{Draw } \tilde{z} \sim N(0, \Delta) \)
9: \( \text{Discretize } \frac{\Delta}{\beta} \text{ by means of (1.74)} \)
10: end for
11: end for

proposed algorithms generate estimates in strict accordance with true prices: the average absolute error is, indeed, lower than 1% for the SIA scheme and 2% for the GVA scheme (where the underperformance of the latter is mainly due to the large error with fewer time steps). The only noticeable inconsistencies concern long-dated (SIA) and short-dated (GVA) out-of-the-money options.

By comparing the accuracy of the schemes, two facts clearly arise: when we consider coarse time grids, the SIA scheme significantly outperforms the GVA scheme. This is particularly true for shorter time horizons (i.e., the 6 months and 1 year case). On the other hand, as we shrink the temporal mesh size, the latter converges faster to the true price. The observed dynamics is well represented in Figure 1.7, where the results for \( T = 1 \) are shown. In the lights of these evidences, the choice of the optimal scheme should be made by taking into account also the designated time framework (both in terms of overall horizon and step width). From a computational point of view, both schemes allow to deal with the discretization of WSVM asset price very efficiently: as shown in Table 1.1, for example, we are able to simulate \( 2 \times 10^5 \) price trajectories with a time step of \( T/100 \) in no more than 21 seconds. With such a time frame we get an error lower than 1% in all but 3 cases. A comparison between the two rows of Table 1.1 also reveals that the SIA scheme is, on average, 20% − 30% slower than the GVA one.

<table>
<thead>
<tr>
<th>Time steps</th>
<th>5</th>
<th>10</th>
<th>50</th>
<th>100</th>
<th>200</th>
<th>300</th>
</tr>
</thead>
<tbody>
<tr>
<td>SIA (s)</td>
<td>1.02</td>
<td>1.96</td>
<td>10.45</td>
<td>21.01</td>
<td>42.24</td>
<td>63.77</td>
</tr>
<tr>
<td>GVA (s)</td>
<td>0.79</td>
<td>1.57</td>
<td>8.50</td>
<td>16.99</td>
<td>35.01</td>
<td>51.39</td>
</tr>
</tbody>
</table>

Table 1.1: Average computational time as function of the number of time steps. The number of simulated paths is fixed to \( 2 \times 10^5 \).

\[12\] All tests have been carried out on the machine illustrated above. Algorithms are written in Matlab code and then compiled as MEX files to achieve better performances.
Chapter 1. Wishart Processes in Finance

It is worthwhile to point out that we only implemented the schemes in their crudest version. Significant improvements can be achieved both in terms of accuracy (i.e. using variance reduction techniques) and in terms of speed (instead of sampling exactly the Wishart process we could use the second or third order schemes developed in the same paper [3] and exploit parallel and GPU computing).

The results outlined in this section look promising: both SIA and GVA schemes turn out to be efficient, fast, and simple to implement. Thanks to the underlying Wishart process sampling [3], moreover, they do not pose any restriction on variance process parameters. A comprehensive comparative analysis with alternative schemes in literature is left to further research.

1.4 The Wishart Affine Stochastic Correlation Model

With the purpose of reproducing well-known multi-asset stylized facts in a tractable way, in [33] the authors introduce the Wishart Affine Stochastic Correlation model (WASC) that makes use of Wishart process to describe the stochastic variance covariance matrix of asset returns. The model proposes the following joint dynamics for a vector of forward asset prices:

\[
df(t) = \text{diag}[f(t)] \sqrt{\Sigma(t)} \, db(t), \quad f(0) \in \mathbb{R}_+^d
\]

where \( \text{diag} [\cdot] \) is the operator that transforms a \( d \)-dimensional column vector into a \( d \times d \) diagonal matrix. In (1.76) \( b(t) \) is a \( d \)-vector Brownian motion such that

\[
b(t) = \sqrt{1 - r^\top r} \, z(t) + W(t) \, r
\]
with $z(t)$ another vector Brownian motion independent on $W(t)$ and $r \in [-1, 1]^d$ such that $r^\top r \leq 1$. Here $r$ can be interpreted as the vector of coefficients meant to drive the linear correlation between the shocks on asset returns and shocks on variance-covariance matrix $\Sigma(t)$. The choice of the correlation structure (1.77) represents the major improvement with respect to the model in [55] and aims at accommodating realistic single asset volatility skews still preserving the affinity of the model. Remarkably, the resulting WASC dynamics (1.76) allows for stochastic correlation among asset returns in a tractable framework where each asset is enriched with a stochastic volatility behavior consistent with the effects observed on plain vanilla markets. The peculiarities of the model can be fully appreciated by referring to the individual, or scalar, dynamics of asset returns $y_i$:

$$
\frac{dy_i(t)}{dt} = -\frac{1}{2} \sum_{j=1}^{d} s_{ij}^2(t) dt + \sum_{j=1}^{d} s_{ij}(t) db_j(t) = -\frac{1}{2} \Sigma_i(t) + \sum_{j=1}^{d} s_{ij}(t) db_j(t), \quad i = 1, ..., d
$$

(1.78)

where $S(t) = S(t)^\top = (s_{ij})_{1 \leq i,j \leq d}$ is the unique positive semi-definite square root of $\Sigma(t)$. We also use, as above, the notation $\Sigma_i(t) = \Sigma_{ii}(t)$ to denote the $i$-th diagonal element of Wishart process. By straightforward computations, from (1.78) we can compute the quadratic covariation of two given assets:

$$
\frac{d[y_k(t), y_l(t)]}{dt} = \frac{d}{dt} \left[ \sum_{j=1}^{d} s_{kj}(t) db_j(t), \sum_{j=1}^{d} s_{lj}(t) db_j(t) \right] = \Sigma_{kk}(t) dt \quad (1.79)
$$

that highlights the role of Wishart process, used to describe the stochastic evolution of the asset returns variance covariance matrix. Furthermore, we can explicitly define the cross-asset correlation matrix $C_Y(t)$ as

$$
C_Y(t) = (\rho_{ij}(t))_{1 \leq i,j \leq d} = \frac{\Sigma_{ij}(t)}{\sqrt{\Sigma_{ii}(t) \Sigma_{jj}(t)}} \quad (1.80)
$$

By exploiting the properties of Wishart process, it can be shown that $C_Y(t)$ is a well-defined correlation matrix (i.e. $C_Y(t)$ is positive semi-definite and each $\rho_{ij}(t) \in [-1, 1]$) as soon as condition (1.7) is satisfied and provided that $\Sigma_{ij}(t) \neq 0$ for $i, j = 1, ..., d$.

Indeed, if $\beta \geq d - 1$, $\Sigma(t)$ is positive semi-definite and $C_Y(t)$ admits the decomposition

$$
C_Y(t) = D^{-1} \Sigma(t) D^{-1} = D^{-1} S(t) S(t)^\top D^{-1} = LL^\top
$$

where $D = \sqrt{\text{diag}[\Sigma(t)]}$. To show that each element $\rho_{ij}(t)$ is bounded in $[-1, 1]$ we use the following theorem that applies for Wishart distributed matrices:

**Theorem 1 (Theorem 2.4.2 in [73])**. Let $X_W \in S_d^+ (\mathbb{R}) \sim \mathcal{W}_d(\beta, \Theta, \Gamma)$, i.e. $X_W$ is a $d \times d$ symmetric matrix that follows a non-central Wishart distribution with degrees of freedom $\beta$, scale $\Theta$ and non-centrality matrix $\Gamma$. Then for a $n \times d$ matrix $B$, we have that $Y = BX_W B^\top \sim \mathcal{W}_n(\beta, B\Theta B^\top, B\Gamma B^\top)$. 

33
Chapter 1. Wishart Processes in Finance

Proof. See [73].

If we set \( X_w = \Sigma(t) \) and \( B = [e_i^d, e_j^d]^\top \) for some admissible \( i \) and \( j \) (with \( e_i^d \) the \( i \)-th element of the standard basis of \( \mathbb{R}^d \)), the previous theorem shows that the resulting matrix \( \tilde{X}_w = \tilde{B}\Sigma(t)\tilde{B}^\top = \begin{bmatrix} \Sigma_i(t) & \Sigma_{ij}(t) \\ \Sigma_{ij}(t) & \Sigma_j(t) \end{bmatrix} \) is a well-defined \( 2 \times 2 \) Wishart process and then it holds that \( \Sigma^2_{ij}(t) \leq \Sigma_i \Sigma_j \).

By combining (1.78), (1.79) with Proposition 3, and fixing a time horizon \( T \), we can represent the \((T\text{-specific})\) \( \text{WASC} \) dynamics as the following \( 2d \times 2d \) system of scalar SDEs (for \( i = 1, \ldots, d \))

\[
dy_i(t) = -\frac{1}{2} \Sigma_i(t) dt + \sqrt{\Sigma_i(t)} dw^i(t),
\]
(1.81)

\[
d\Sigma_i(t) = \kappa_i(\theta_i - \Sigma_i(t)) dt + \eta_i \sqrt{\Sigma_i(t)} dw_i(t)
\]
(1.82)

where the parameters \( \kappa_i, \theta_i \) and \( \eta_i \) are given, respectively, in (1.16), (1.17) and (1.18).

Here the correlation structure among Brownian motions \( w = [w^1, w^2, \ldots, w^d, w^1, w^2, \ldots, w^d]^\top \) is described by means of the stochastic block matrix

\[
C(t) = \begin{bmatrix}
C_{\Sigma}(t) & C_{\Sigma\Sigma}(t) \\
C_{\Sigma\Sigma}(t) & C_{\Sigma}(t)
\end{bmatrix}
\]
(1.83)

where the submatrices, other than \( C_{\Sigma} \) already introduced in (1.80), will be described in the following. Interestingly, from (1.81) and (1.82) we have that, for any \( T > 0 \), the scalar dynamics of each asset is consistent with a standard Heston model driven by the \( i \)-th diagonal element of \( \Sigma(t) \). This assures that the behaviour induced by \( \text{WASC} \) in terms of reconstructed implied volatility surfaces is in line with the documented findings of traditional one factor stochastic volatility models. Consistently with Heston model, the asset specific returns-volatility correlation is constant: following [33], we have

\[
\text{Corr}_t [dy_i(t), d\Sigma_j(t)] = \rho_{ij} dt = \frac{\text{Tr}[QR]}{\sqrt{Q_{ii}}} dt
\]
(1.84)

where \( R_i \) is the matrix with \( r \) on the \( i \)-th row and zero elsewhere. We can even generalize the previous result by explicitly deriving the correlation between the \( i \)-th log-asset and a generic diagonal element of \( \Sigma(t) \): we, indeed, have that (as shown in [33]) the covariation between asset returns and volatility terms is given by

\[
d [y_i(t), \Sigma_j(t)] = 2\Sigma_{ij}(t) \text{Tr}[R_j Q] dt.
\]
(1.85)

By combining (1.85) and (1.79) and exploiting (1.4), we get the generic element of matrix \( C_{\Sigma\Sigma}(t) \) as

\[
\text{Corr}_t [dy_i(t), d\Sigma_j(t)] = \frac{2\Sigma_{ij} \text{Tr}[R_j Q]}{\sqrt{\Sigma_i} \sqrt{4\Sigma_j Q_{jj}}} dt = \frac{\text{Tr}[R_j Q]}{\sqrt{Q_{jj}}} \frac{\Sigma_{ij}}{\sqrt{\Sigma_i} \sqrt{\Sigma_j}} dt = \rho_{ij}(\rho_{ij}(t)) dt.
\]
(1.86)
In the last equality of (1.86), we introduce a new representation of such correlations that can be seen, quite fascinatingly, as the product of the (constant) proper, or scalar, $j$-th asset-volatility correlation and the cross-asset correlation between $i$-th and $j$-th assets. This result highlights the peculiar dependence structure inherent in the WASC model and could help in gaining more insights on parameters impact on correlation surfaces in the spirit of the study carried out in [33]. Further, using (1.4), it follows that the elements of $\Sigma(t)$ have the form

$$\text{Corr}_t [d\Sigma_i(t), d\Sigma_j(t)] = \frac{4\Sigma_{ij} Q_{ij}}{4\sqrt{\Sigma_i\Sigma_j\Sigma_{ii}Q_{jj}}} dt = \frac{Q_{ij}^i}{\sqrt{Q_{ii}^i Q_{jj}^j}} \rho_{ij}(t) dt. \quad (1.87)$$

From (1.80), (1.86) and (1.87), we have that the stochastic evolution of (1.83) is fully described by processes $\rho_{ij}(t)$. Let us consider, for example, the case $d = 2$: matrix $C(t)$ then reads as

$$C(t) = \begin{bmatrix}
1 & q_{12} \rho_{12}(t) & \rho_1 & \rho_1 \rho_{12}(t) \\
q_{12} \rho_{12}(t) & 1 & \rho_2 & \rho_2 \\
\rho_1 & \rho_2 & 1 & \rho_{12}(t) \\
\rho_1 \rho_{12}(t) & \rho_2 & \rho_{12}(t) & 1
\end{bmatrix} \quad (1.88)$$

where $q_{ij} = \frac{Q_{ij}^i}{\sqrt{Q_{ii}^i Q_{jj}^j}}$. This representation highlights the peculiar dependence structure induced in the WASC model and could provide useful insights on the role of $Q$ and $r$ in determining the relation among state variables.

### 1.4.1 A restricted version of the model

In this section we consider a restricted, more intuitive, specification of WASC model: we assume matrix $M$ to be diagonal and with negative entries. This setting leads to a very interesting dynamics for the diagonal elements of Wishart process. From Proposition 3, if $M$ is diagonal, we have

$$\kappa_i = -2M_{ii}, \quad (1.89)$$

$$\eta_i = 2\sqrt{Q_{ii}^i}, \quad (1.90)$$

for all $T > 0$. The proof of this result (which follows by direct computation) is left to the interested reader. An immediate consequence is that the asset instantaneous variances are now described by time-independent CIR processes, in the sense that the parameters involved are no longer function of the time horizon considered.

In our opinion the resulting parametrization turns out to be the most genuine multi-asset extension of the Heston model: each asset is exactly described by a single instance of the Heston dynamics while the joint behaviour is enriched by cross-assets and cross-variances stochastic correlation, all wrapped in an affine framework. As far as we
know, there are no alternative settings that can reach a comparable degree of flexibility. The exact Heston representation of asset dynamics also helps in understanding the role and the impact of WASC parameters, that in the general formulation appear somehow unclear. In particular, it is worthwhile to point out that the pricing of single-asset European claims is only affected by the corresponding diagonal element of $\Sigma_0$. To see this, it suffices to notice that for $\lambda = \lambda_i e_i^d$, the matrix $A_y(\tau)$ in (1.94) has a non-null $i$-th diagonal element and zeros elsewhere. This peculiarity is in line with the asymptotic analysis provided in [32] where the $i$-th implied volatility approximation for short time to maturity is found to be

$$\sigma_{imp,i}^2 = \Sigma_{ii} + (r_1 Q_{1i} + r_2 Q_{2i}) m f + \frac{1}{2} \left( \frac{Q_1^2 + Q_2^2}{2\Sigma_{ii}} - \frac{7(r_1 Q_{1i} + r_2 Q_{2i})^2}{6\Sigma_{ii}} \right) m f^2, \quad (1.91)$$

where $m f = \log \left( \frac{K}{f_i(0)} \right)$ denotes the log-forward moneyness. Consequently, the off-diagonal entries of $\Sigma_0$ can be used to match multi-asset stylized facts without compromising the shape of individual volatility surfaces. This represents and additional degree of freedom that in the general WASC model we would not have. A possible calibration strategy could be to set the off-diagonal entries of $\Sigma_0$ in order to match a predefined initial cross-asset correlation matrix. Alternatively, provided that liquid multi-asset derivatives are traded, we could try to fit the implied correlation market evities. In order to further develop this point, we now study the impact of $\Sigma_{12}$ on the price of Best-Of put options whose payoff is

$$(K - \max [S_1(T), S_2(T)])^+. \quad (1.92)$$

For the numerical analysis we assume the following WASC parameters:

$$\beta = 1.1, \Sigma_0 = \begin{bmatrix} 0.04 & 0 \\ 0 & 0.04 \end{bmatrix}, \quad M = \begin{bmatrix} -0.7 & 0 \\ 0 & -1.2 \end{bmatrix}, \quad Q = \begin{bmatrix} 0.3 & 0.3 \\ 0.2 & 0.3 \end{bmatrix}, \quad r = \begin{bmatrix} -0.6 \\ -0.1 \end{bmatrix}$$

that are meant to describe realistic market scenarios and allowing for a well defined Wishart process. Figure [1.8] shows the implied correlation profiles corresponding to different values of $\Sigma_{12}$, where the implied correlation is defined to be the value of parameter $\rho$ such that the WASC price equals the one obtained in a 2 assets Black-Scholes setting, i.e.

$$P_{Best-Of}^{BS}(\sigma_{imp,1}^{WA}(K, T), \sigma_{imp,2}^{WA}(K, T), \rho, K, T) = P_{Best-Of}^{WA}(\sigma_{imp,1}^{WA}(K, T), K, T). \quad (1.93)$$

Here $\sigma_{imp,i}^{WA}(K, T)$ is the Black-Scholes implied volatility corresponding to the option written on the $i$-th asset with strike $K$ and maturity $T$ whose price is computed with WASC model. By exploiting the affinity of the 2 model, Best-Of put options are priced by numerically computing a bi-dimensional inverse Fourier transform. We refer the

---

Without loss of generality only the case $d = 2$ is considered.
1.4. The Wishart Affine Stochastic Correlation Model

Figure 1.8: Best-Of put options implied correlation skew for different values of $\Sigma_{12}$. Other parameters used: $f_1(0) = f_2(0) = 100$, $T = 1$ and $r = 0\%$.

reader to [33] where this pricing methodology is developed for Best-Of contracts. From the numerical results, it is evident that the off-diagonal Wishart element plays a significant role in modelling the implied correlation skew: we observe an increase in implied correlation levels for higher values of $\Sigma_{1,2}$. This, in turn, induces an increase in option prices consistently with the fact that Best-Of put options are long correlation products that benefit from lower assets returns dispersion.

1.4.2 WASC Characteristic function

As in the WSVM case, the chosen correlation structure (1.77) assures the affinity of WASC model. This means, once more, that we can express the (joint) characteristic function of the asset returns vector $y(t)$ as an exponential affine transformation of state variables $y(t)$ and $\Sigma(t)$ as recalled in the following Proposition:

**Proposition 7 (Joint characteristic function of log-prices in WASC).** Let the log-forward prices vector $y(t)$ be described by (1.78) and $\lambda$ be an auxiliary vector-valued variable $\lambda = [\lambda_1, ..., \lambda_d]^\top$. Then for $T > t$, the WASC (conditional) characteristic function of $y(T)$ admits the following closed formula representation

$$
\phi_{yA}(\lambda, \tau) = \mathbb{E} \left[ \exp \left( i \langle \lambda, y(T) \rangle \right) | y(t) \right] = \exp \left( i \langle \lambda, y(t) \rangle + \text{Tr} [A_{y}(\tau)\Sigma(t)] + b_{y}(\tau) \right),
$$

(1.94)

with the deterministic matrix $A_y(\tau)$ and the scalar function $b_y(\tau)$ given by

$$
A_y(\tau) = A_{22}(\tau)^{-1} A_{21}(\tau),
$$

$$
b_y(\tau) = -\frac{\beta}{2} \text{Tr} \left[ \log (A_{22}(\tau)) + \tau (M + \iota \lambda Q^\top R^\top) \right],
$$

37
Chapter 1. Wishart Processes in Finance

and

\[
\begin{bmatrix}
A_{11}(\tau) & A_{12}(\tau) \\
A_{21}(\tau) & A_{22}(\tau)
\end{bmatrix} = \exp \left( \tau \begin{bmatrix}
M + iQ^T r \lambda^T & -2Q^T Q \\
-\frac{1}{2} (\lambda \lambda^T + i \text{diag} [\lambda]) & -(M + iQ^T r \lambda^T)^T
\end{bmatrix} \right).
\]

Proof. See [33].

Thanks to Proposition 7, we are able to price both plain vanilla and multi-asset options (if transform-based techniques are applicable\(^{14}\)) in a comprehensive framework. In particular, we price options on the \(i\)-th asset as a basket option with degenerate weights vector \(e_i^d\), such that \(\lambda = \lambda_i e_i^d\). Furthermore, the knowledge of the joint characteristic function of asset returns vector allows to make use of bounds techniques as those developed in [19] for basket options. Despite the analytical tractability, several numerical issues arise when we try to calibrate WASC model to market data by exploiting (1.94). Not only, indeed, we have to evaluate functions of matrix argument for each computation of the characteristic function (as in the WSVM case), but, even worse, we are required to perform \(d\) different plain vanilla pricing (one for each asset) for a single parameters set. This is due to the lack of liquid multi-asset derivatives that force us to calibrate model parameters to the individual market implied volatility surfaces. As reported in [32], such a naive algorithm can take up to 15 minutes in the simplest case \(d = 2\). This is, clearly, not feasible for real market applications.

1.4.3 A new calibration procedure

In this section we present an innovative and efficient methodology to calibrate WASC parameters that exploits the close link existing between Heston model and marginal WASC dynamics. The proposed algorithm is firstly tested in a simplified framework and then applied to market data. The results obtained also highlight the impact of parameter \(\beta\) on model accuracy in reproducing market volatility smiles.

Let us consider a WASC parameters set \(\pi_{WA}\) (with cardinality \(N_{WA}\)) and fix a maturity \(T\). For the generic \(i\)-th asset described by (1.78) we can define a function \(g_i^{H-WA}\) that maps WASC parameters to those of a scalar Heston dynamics. In other words, we set \(g_i^{H-WA} : \mathbb{R}^{N_{WA}} \times \mathbb{R}_{>0} \rightarrow \mathbb{R}^5\) such that \(g_i^{H-WA}(\pi_{WA}, T) = \pi_{i,H} = [v_{0,i}, \kappa_i, \theta_i, \eta_i, \rho_i]^T\) as defined in (1.15)-(1.18) along with the assets-volatility correlation (1.84). Consequently, for calibration purposes, we can replace the cumbersome WASC characteristic function with the simpler Heston one. Furthermore, we can compute analytically the gradient of the objective function with respect to WASC parameters. In Appendix 2.A we show how to compute explicitly the matrix

\[J_{i,T}^{H-WA}(\pi_{WA}) = \]

\(^{14}\)We refer here to those cases in which transformed payoff functions are available and evaluation of (inverse) transforms are numerically feasible.
\[ \nabla g_i^{H-W_A} \in \mathbb{R}^{N_{WA} \times 5}, \text{i.e. the Jacobian matrix of function } h_i^{WA} \text{ with elements:} \]
\[ J_{H-W_A}^{j,r} = \frac{\partial g_i^{H-W_A} \left( \pi_{WA}, T \right)}{\partial \pi_{WA,q}}. \]  
(1.95)

Then, the Jacobian matrix of \( \tilde{r}_{i,T}(\pi_{WA}) \) (the residuals vector composed of options with maturity \( T \) written on the \( i \)-th asset) can be written as
\[ J_{WA}^{i,T}(\pi_{WA}) = J_{H-W_A}^{i,T}(\pi_{WA}) J_i^H \left( g_i^{H-W_A}(\pi_{WA}, T) \right) \]  
(1.96)

where, once more, the second matrix in the right-hand side of (1.96) is known thanks to [30]. To obtain the overall Jacobian matrix \( J_{WA} \), we simply need to compute (1.96) for each maturity and asset taken into account and aggregate the resulting matrices. As illustrated for the single-asset case, finally, the gradient of objective function (1.54) is given by
\[ \nabla f_{obj} = J_{WA}^{i,T} \tilde{r}(\pi_{WA}). \]

The calibration algorithm so defined avoids the computation of WASC characteristic function and significantly reduces the issues due to the possible presence of multiple minima. Given that we rely on the law identity in Proposition 3 rather than on some approximation, the routine does not require any further step.

A simplified calibration exercise

The accuracy of the proposed algorithm is illustrated by considering the following numerical experiment: let us suppose that a fictitious two-assets market is perfectly described by the WASC parameters reported in the first column of Table 1.6. Even if simplified, the data outline realistic market environments: they represent the calibrated parameters set (truncated at the first significant decimal digit) found in [32] for the couple of indices EuroStoxx50-DAX. We construct a full implied volatility surface for each asset assuming to have options with maturities \( T = [0.25, 0.5, 1, 3] \) and 41 equally spaced strikes ranging from 0.5 to 1.5 (initial asset values are set for simplicity equal to 1). Each of the resulting surfaces consists of 164 options. The goal is to implement the proposed algorithm in order to find a suitable parameters set that reproduces the supposed market data. Hopefully, we expect the calibrated parameters to be reasonably close to the original ones (accuracy) and to experience a limited dependency on the initial guess (robustness). For this test we set the starting values of the optimization routine as shown in the second column of Table 1.6. The choice is meant to assess the robustness of the algorithm in the case in which the initial guess is very far from the optimal set. Indeed, not only the discrepancy is mixed - some values are overestimated, others underestimated - but the distance between initial guess and optimal values is substantial: the smallest gap, defined as percentage difference, is equal to 35.29%. The mistaken initialization of the problem and the high dimensionality of the parameters space make the calibration task more challenging and could potentially lead to suboptimal outcomes. Notwithstanding, the proposed algorithm is able to produce results very
close to the original values: the norm of the errors between true prices and calibrated ones is \(2.2069 \times 10^{-7}\). Most remarkably, the procedure takes only 3.56 seconds on a consumer laptop PC with an Intel Core i7 CPU and 8 GB RAM. By considering parallelization and porting to more efficient languages we can obtain a further speedup. It is worthwhile also noting that in realistic applications, the calibration problem is somehow facilitated thanks to the availability of previous optimal sets that act as efficient guesses. In the lights of all these evidences, we believe that the proposed methodology represents a highly efficient tool for the calibration of WASC model. This is particularly true if we intend to increase the number of assets involved with the subsequent growth of dimensionality.

### Calibration to market data

We now want to validate the procedure with realistic market data. Despite the general applicability of the algorithm, we focus our attention on the reduced specification of the model. With this in mind, we select a basket of market quoted instruments composed of 201 European call options written on EuroStoxx50 index and 182 on DAX index. The set of derivatives on the DAX is the same set used to calibrate the WSVM model. We further set, for simplicity, interest rates and dividends to zero. Thanks to the efficiency of the new calibration algorithm, we are able to calibrate model parameters in less than 3 seconds. The outputs of the optimization routine are shown in the leftmost column of Table 1.7 in Appendix I.C.2. Given that, as illustrated above, the off-diagonal element of \(\Sigma_0\) does not impact the pricing of univariate call options, we set its value such that the initial correlation among the two indices equal the one-year historical one (that is found to be 0.9715).

The most interesting result is that \(\beta\) is lower than 1. This is coherent with the evidences in [32] where similar results are found. Figure 1.9 shows the calibrated model implied volatility skews for the two indices with respect to maturities of one month, one year and three years. The model succeeds in reproducing the shape of market volatility surfaces but the mispricing is not negligible for short term far-from-the-money options. This is particularly true for the EuroStoxx50 index as highlighted from the fact that the error in volatility terms is roughly 3 times higher than the error made for the DAX. Additionally, we can compare the evidences from the calibration of the two Wishart based models against the results for the Heston and Bi-Heston models calibrated to the same basket of DAX options. Table 1.2 shows the calibrated initial variance of asset returns along with the Mean Squared Error with respect to both price and implied volatility for the four models. Consistently, the estimates of initial variance are in strict agreement: all the models agree on the initial volatility.

The two multi-factor models tend to perform quite similarly (although errors for WSVM are slightly smaller) and substantially outperform the simpler Heston and WASC dy-

---

\(^{15}\)We simply consider the historical value as a target correlation level defined by model user.
1.4. The Wishart Affine Stochastic Correlation Model

Figure 1.9: Calibration results for WASC. Resulting value for parameter $\beta$ is 0.8577. Comparison with market implied volatility for EuroStoxx50 (left) and DAX (right) indices for selected tenors.

In particular, by comparing the error of WSVM and WASC models, the out-performance of the former is clearly evident. This is not surprisingly since we contrast a multi-factor volatility setting (WSVM) with the WASC single-asset dynamics that, as developed in Section 1.4, is equivalent to 1-factor parametrizations. It is important to remark, however, that the models are meant to address rather different tasks (i.e. single-asset and multi-asset modelling).

<table>
<thead>
<tr>
<th></th>
<th>WSVM</th>
<th>Bi-Heston</th>
<th>WASC</th>
<th>Heston</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial Variance</td>
<td>(\text{Tr} [\Sigma_0])</td>
<td>(v_{0,1} + v_{0,2})</td>
<td>(\Sigma_{12})</td>
<td>(v_0)</td>
</tr>
<tr>
<td>Calibrated Value</td>
<td>0.0866</td>
<td>0.0868</td>
<td>0.0890</td>
<td>0.0842</td>
</tr>
<tr>
<td>Error Price</td>
<td>2.45E-06</td>
<td>2.83E-06</td>
<td>2.97E-06</td>
<td>2.95E-06</td>
</tr>
<tr>
<td>Error Vol</td>
<td>2.85E-05</td>
<td>3.56E-05</td>
<td>9.88E-05</td>
<td>1.07E-04</td>
</tr>
</tbody>
</table>

Table 1.2: Comparison of calibration outputs on DAX index. The Initial Variance row specifies the initial variance of asset returns in the corresponding model. Complete calibration results are given in Table 1.4 for WSVM, Table 1.5 for Bi-Heston, Table 1.7 for WASC and Table 1.8 for Heston model.

Notwithstanding the differences between the 2 models, once more, some fix is required in order to enforce the existence and uniqueness condition for matrix process \(\Sigma(t)\). As done in the single-asset framework, we tackle the calibration problem imposing \(\beta \geq 1\). Results are exhibited in the second column of Table 1.7. Even if the loss in accuracy seems to be somehow limited (the error measure are just slightly higher than...
in the unconstrained setting), a relevant issue arises: in Figure 1.10 we report simulated trajectories of WASC cross-asset correlation obtained with the resulting parameters set. The fact that $\beta$ satisfies condition (1.7) effectively ensures $\rho_{12}(t)$ to lie in the range $[-1, 1]$. However, the fact that the parameter is just slightly above the threshold ($\beta = 1$) makes the boundary of $\mathcal{S}^+_2(\mathbb{R})$ very likely to be attained. Very often, then the absolute value of correlation is stuck at 1. We can also experience sudden changes in correlation from $+1$ to $-1$ (or vice versa) in a very restricted time frame (even on a daily basis).

In order to study the dependence of the observed phenomenon on the initial value of correlation, in the rightmost panel of Figure 1.10 we also consider the case $\Sigma_{12} = 0$. 

**Figure 1.10:** Simulated trajectories of cross-asset correlation $\rho_{12}(t)$ for $t \in [0, 1]$ generated with calibrated parameters obtained imposing $\beta \geq 1$. Left panel: $\Sigma_{12} = 0.0654$. Right panel: $\Sigma_{12} = 0$.

**Figure 1.11:** Simulated trajectories of cross-asset correlation $\rho_{12}(t)$ for $t \in [0, 1]$ generated with calibrated parameters obtained imposing $\beta \geq 3$. Left panel: $\Sigma_{12} = 0.0643$. Right panel: $\Sigma_{12} = 0$. 

that produces a similar erratic correlation dynamics.

Given the intent to apply WASC model to describe the joint behaviour of asset prices, this represents a major issue. To tackle the problem, we decide to enforce the positive definiteness condition for $\Sigma(t)$, given by (1.8), that in our setting equals to set $\beta \geq 3$. Calibrated parameters are collected in the rightmost column of Table 1.7. The corresponding cross-asset correlation dynamics is depicted in Figure 1.11: the trajectories are now much more meaningful. Further, as a consequence of the fact that $\Sigma(t)$ is defined on the interior of $S^+_2(\mathbb{R})$, $\rho_{1,2}(t)$ is bounded in $(-1, +1)$. Nonetheless, the stronger condition enforced has a severe impact on the ability of the model to reproduce single-asset market evidences. Reconstructed volatility skews are shown in Figure 1.12. Significant discrepancies now emerge for far-from-the-money options. In particular, in the very short-end of the volatility term structure the error with respect to market volatilities can be as high as 11.87% (in-the-money options on EuroStoxx50) and 10.80% (out-of-the-money options on DAX). Disappointingly, we face a non trivial trade-off between plain vanilla pricing accuracy and realistic modelling of cross-asset correlation. A possible solution to mitigate the problem could be to set $\beta$ equal to some value in the range $(1, 3]$. This alternative, however, would require to couple the plain vanilla analysis with adequate market evidences on multi-asset derivatives. The topic will be further developed in Chapter 2 and represents the main motivation for the introduction of a new Wishart-based multi-asset pricing model.

1.4.4 A new simulation scheme for the WASC

This section is devoted to present the first simulation algorithm specifically devised for WASC model. As far as we know, indeed, there are no previous attempts in literature
Chapter 1. Wishart Processes in Finance

to deal with the discretization of prices trajectories (1.76). In particular, our task is to develop an efficient, yet accurate, scheme to discretize the system of SDEs (1.81)-(1.82). It is evident that a standard discretization (e.g. via Euler scheme) is unfeasible, since we also need to take into account the evolution of non diagonal elements of $\Sigma(t)$ to determine the dependence structure and satisfy the positive semi-definiteness constraint for the Wishart process. We overcome such limitations by considering the Wishart process sampling scheme developed in [3]: that is, we consider as given an entire discretized path of $\Sigma(t)$ over the time grid $0 = t_0 < t_1 < \cdots < t_M = T$ with time step $\Delta$. In this way we are only left with the problem of sampling the log-prices trajectories. The most challenging task here is to embed the correlation structure (1.83) in the discretization of (1.81). In standard cases, we would compute the Cholesky decomposition of matrix $C(t)$ such that

$$[w_1^\Sigma, w_2^\Sigma, ..., w_d^\Sigma, w_1^\nu, w_2^\nu, ..., w_d^\nu]^\top = LC(t) \left[ w_1^*, w_2^*, ..., w_d^*, w_{d+1}^*, w_{d+2}^*, ..., w_{2d}^* \right]^\top$$

(1.97)

where $L_C(t) = (\ell_{i,j}(t))_{1 \leq i \leq d, 1 \leq j \leq i}$ is the lower triangular matrix that satisfies $C(t) = L_C(t)L_C(t)^\top$ and $w^* = [w_1^*, w_2^*, ..., w_d^*, w_{d+1}^*, w_{d+2}^*, ..., w_{2d}^*]^\top$ is a vector of independent Brownian motions. By exploiting (1.97), we can rewrite (1.81) as

$$dy_i(t) = -\frac{1}{2}\Sigma_i(t)dt + \sqrt{\Sigma_i(t)} \sum_{j=1}^{d+i} \ell_{d+i,j}(t)dw_j^*(t)$$

(1.98)

that can be discretized by generating $2d$ independent gaussian random variables. Unfortunately, as already stressed in section [1.3.3] in our setting this is not readily doable as a consequence of the mechanics of Wishart sampling algorithm. In other words, we do not have a direct “access” to the discretized paths of $w_1^\Sigma, w_2^\Sigma, ..., w_d^\Sigma$.

The simple idea underlying the new simulation scheme is to exploit the auxiliary scalar dynamics of $\Sigma_i(t)$ to get an approximation of $w_i^\Sigma(t+\Delta) - w_i^\Sigma(t)$. Let $\tilde{\Sigma}(t)$ and $\Sigma(t+\Delta)$ be the realizations of the trajectory of Wishart process for 2 adjacent points on the time grid computed by means of the exact scheme in [3]. The discretized version of (1.1), that reads

$$\tilde{\Sigma}(t+\Delta) - \tilde{\Sigma}(t) \approx \kappa_i \left( \theta_i - \tilde{\Sigma}(t) \right) \Delta + \eta_i \sqrt{\Sigma(t)} \tilde{w}_i$$

(1.99)

can now be used to approximate the gaussian variable $\tilde{w}_i$. Let $\tilde{w}_\Sigma$ be the result of (1.99), for a sufficiently small time interval, $\tilde{w}_\Sigma = [\tilde{w}_\Sigma_1, \tilde{w}_\Sigma_2, ..., \tilde{w}_\Sigma_d]^\top$ represents an approximation of a vector of gaussian variables with correlation matrix $C_\Sigma(t)$ (i.e. the realization at time $t$ of matrix $C_\Sigma$). Further, let $\hat{L}_\Sigma(t)$ be the lower triangular matrix obtained from the Cholesky decomposition of $C_\Sigma(t)$, then

$$\tilde{w}_\Sigma^* = \hat{L}_\Sigma^{-1}(t)\tilde{w}_\Sigma$$

(1.100)

is composed of $d$ approximated independent gaussian random variables. By sampling an additional random vector $\tilde{w}_\gamma^*$ from $N(0_d, \Delta I_d)$ (the $d$-variate gaussian distribution)
and setting
\[ \hat{\mathbf{w}}^* = \begin{bmatrix} \hat{\mathbf{w}}_{\Sigma}^* \\ \hat{\mathbf{w}}_y^* \end{bmatrix} \] (1.101)
we can finally approximate (1.98) as
\[ \hat{y}_i(t+\Delta) = \hat{y}_i(t) - \frac{1}{2} \hat{\Sigma}_i(t) \Delta + \sqrt{\hat{\Sigma}_i(t)} \sum_{j=1}^{d+i} \hat{\ell}_{d+i,j}(t) \hat{w}_j^*(t) \] (1.102)

where \( \hat{L}_C(t) \) results from the factorization of \( \hat{C}(t) \). If \( \hat{C}(t) \) turns out not to be positive definite, we take its positive part, \( \hat{C}^+(t) \), (defined as the matrix obtained from the spectral decomposition of \( \hat{C}(t) \) with negative eigenvalues replaced by zeros) and apply the extended Cholesky decomposition described in [54]. The complete algorithm is exhibited in Algorithm 3. In full analogy with the WSVM case, we refer to the new scheme as the Gaussian Variables Approximation (GVA) scheme for WASC model.

Even if this new scheme would apply to the general specification of the model, here we suppose to deal with the reduced model presented in Section 1.4.1 (i.e. we consider matrix \( M \) to be diagonal). In particular, we develop an extensive numerical investigation based on the parameters set calibrated to market data enforcing the condition \( \beta \geq 3 \) and shown in the rightmost column of Table 1.7. Considering \( 5 \times 10^5 \) simulation paths, we price European call options written on any of the 2 assets with maturity \( T = 1 \) and moneyness in the range \{70\%, 100\%, 130\%\}. For the sake of simplicity, we assume interest rates and dividends equal to zero and \( f_1(0) = f_2(0) = 100 \). Further, we extend the analysis to the pricing of multi-asset contracts in order to validate the ability of the new scheme to capture the (rather) involved dependence structure induced by the WASC model. Given the availability of (semi-)analytical pricing, we convey our attention to the Best-Of put options described above.

As comparison, we implement an adapted version of the scalar full truncated Euler (TE) scheme that reads
\[ \hat{y}(t+\Delta) = -\frac{1}{2} \text{Vec}\left[ \hat{\Sigma}^+(t) \right] \Delta + \sqrt{\hat{\Sigma}^+(t)} \left( \sqrt{1 - \mathbf{r}^\top \mathbf{r}} \, \hat{\mathbf{z}} + \hat{\mathbf{W}} \, \mathbf{r} \right) \sqrt{\Delta} \] (1.103)
\[ \hat{\Sigma}(t+\Delta) = \hat{\Sigma}(t) + \left( \beta \mathbf{Q}^\top \mathbf{Q} + M \hat{\Sigma}^+(t) + \hat{\Sigma}^+(t) M^\top \right) \Delta \] (1.104)
\[ + \sqrt{\hat{\Sigma}^+(t)} \hat{\mathbf{W}} \mathbf{Q} \sqrt{\Delta} + \mathbf{Q}^\top \hat{\mathbf{W}}^\top \sqrt{\hat{\Sigma}^+(t)} \sqrt{\Delta} \] (1.105)

with \( \hat{\mathbf{z}} \) and \( \hat{\mathbf{W}} \), respectively, \( d \)-dimensional vector and square matrix of independent standard gaussian random variables. Here Vec \([ \cdot ]\) is the operator that extracts the elements on the main diagonal of a square matrix into a column vector.

In the context of plain vanilla options, results in Tables 1.15-1.18 show that both schemes allow for accurate price estimates as the size of the time step is sufficiently small. More in detail, the TE scheme is found to outperform the GVA scheme when
Chapter 1. Wishart Processes in Finance

Algorithm 3 The Gaussian Variables Approximation (GVA) scheme for the WASC

1: for each asset $i$, $i = 1, ..., d$ do
2:   Compute parameters $\kappa_i$, $\theta_i$, $\eta_i$ and $\rho_i$ as given in Proposition 3
3: end for
4: for each simulation trial $n$, $n = 1, ..., N$ do
5:   Initialize $\hat{\Sigma}(0) = \Sigma_0$
6:   Initialize $\hat{y}(0) = \log(f(0))$
7:   for each time-step $t_m$, $m = 0, ..., M - 1$ do
8:      Compute correlation matrix $\hat{C}(t_m)$ as given by (1.83)
9:      Compute $\hat{L}_C(t_m)$ from Cholesky decomposition of $\hat{C}(t_m)$ or, in alternative, from the extended Cholesky decomposition of $\hat{C}^+(t)$
10:     Sample the Wishart process for a time step of $\Delta$ with initial state $\hat{\Sigma}(t_m)$ using the scheme in [3]
11:     Approximate the $d$ gaussian random variables $\tilde{w}$ by means of (1.99)
12:     Compute $\tilde{w}^*_\Sigma = \hat{L}^{-1}_\Sigma(t)\tilde{w}_\Sigma$
13:     Sample $\tilde{w}^*_\Sigma$ from $N(0_d, \Delta I_d)$
14:     Discretize (1.98) by means of (1.102)
15:   end for
16: end for

The time grid is coarse (10, 20 and 50 steps per year), while for smaller mesh widths the 2 approaches tend to perform similarly. In particular, if we consider only the case of 200 time steps per year, the GVA scheme outperforms the TE scheme in 4 instances out of 6. With this setting and taking into account option prices for both assets, the absolute mean percentage error is respectively equal to 0.232% for the GVA scheme and to 0.275% for the TE scheme. It is worthwhile to remark, though, that the GVA scheme systematically requires a finer time discretization to produce reliable estimates (true prices lying in the 95% confidence interval) compared to the simpler TE scheme. This is due to the fact that the approximation exploited in (1.99) seems to be adequate only for very small time intervals. When it comes to the pricing of multi-asset options, quite surprisingly, the TE scheme seems to produce more accurate estimates: the percentage error reported in Table 1.19 is systematically lower than 1%. On the other hand, results for the GVA scheme in Table 1.20 highlight some inconsistencies in the pricing of the deep out-of-the money option.

From a computational point of view, the TE scheme greatly outperforms the GVA scheme with the latter that results $110\% - 130\%$ slower than the former as exhibited in Table 1.3. When implementing the GVA scheme, indeed, at each time step we are asked to perform the Cholesky factorization of the $2d$-dimensional matrix $\hat{C}(t)$ with a considerable increase in the computational burden.

Nonetheless, the new scheme proposed embeds the inherent advantage to implement the exact sampling of Wishart process thanks to the algorithm in [3]. This feature is
1.5. Concluding Remarks

<table>
<thead>
<tr>
<th>Time steps</th>
<th>10</th>
<th>20</th>
<th>50</th>
<th>100</th>
<th>200</th>
</tr>
</thead>
<tbody>
<tr>
<td>TE (s)</td>
<td>5.48</td>
<td>10.74</td>
<td>25.84</td>
<td>53.56</td>
<td>104.08</td>
</tr>
<tr>
<td>GVA (s)</td>
<td>11.66</td>
<td>23.48</td>
<td>59.68</td>
<td>119.74</td>
<td>244.54</td>
</tr>
</tbody>
</table>

Table 1.3: Computational time as function of the number of time steps. The number of simulated paths is fixed to $5 \times 10^5$. Results refer to the implementation on a laptop PC with an Intel Core i7 CPU and 8 GB RAM via Matlab MEX files.

of great importance in all the cases in which we need to estimate the (conditional) moments of the distribution of the elements of $\Sigma(t)$. A relevant example of such a task will be presented in the next chapter in conjunction with the calibration of a new pricing model that can be seen as a generalization of (1.76). The ability to consistently deal with the discretization of Wishart process is, indeed, the main reason that led us to develop the new scheme.

### 1.5 Concluding Remarks

The matrix structure of Wishart-based stochastic volatility models provides a remarkable degree of flexibility in describing the evolution of asset(s) volatility. Realistic implementations, though, require the development of specific numerical techniques in order to deal with the inherent level of complexity. In this Chapter we have shown, leveraging on a thorough analysis of distributional properties of Wishart process, some possible solutions intended to make this class of model more suitable for real market applications. Accordingly, we hope that our contribution will increase the interest of researchers and practitioners towards matrix-variate stochastic volatility dynamics. It is important, however, to point out that existing Wishart-based models suffer from a significant inconsistency in the pricing of European claims when conditions for existence and uniqueness of Wishart process are enforced. This remark is the crucial argument that leads to the introduction of a more general class of pricing models discussed in Chapter 2.
Appendices to Chapter 1

1.A Heston and Bi-Heston approximation of WSVM

In this section we present the auxiliary results needed to obtain the (semi-)analytic formulation of the gradient of call option prices when the WSVM is approximated by the Heston or the Bi-Heston model.

1.A.1 The Heston model case: the approach of Cui et al. [30]

The characteristic function of log-asset in the Heston model with parameters set \( \pi_H = [v_0, \kappa, \theta, \sigma, \rho]^T \) is known to be (as originally derived in [63])

\[
\phi_y^H(\lambda, T) = \exp \left( i \lambda y(0) + \frac{\kappa \theta}{\sigma^2} \left( (e_H + d_H) T - 2 \log \left( \frac{1 - l_H e^{d_H T}}{1 - l_H} \right) \right) + \frac{v_0}{\sigma^2} (e_H + d_H) \frac{1 - e^{d_H T}}{1 - l_H e^{d_H T}} \right)
\]  

(1.106)

where we set, for simplicity, interest rates and dividends to zero. Further, we have

\[ y(0) = \log (f(0)) \]

and

\[
e_H = \kappa - \sigma \rho \mu \lambda, \\
d_H = \sqrt{e_H^2 + \sigma^2 (\lambda^2 + i \lambda)} \\
l_H = \frac{e_H + d_H}{e_H - d_H}
\]

(1.107)  (1.108)  (1.109)

In [30], the authors show that (1.106) can be written as

\[
\phi_y^H(\lambda, T) = \exp \left( i \lambda y(0) - \frac{\kappa \theta \rho}{\sigma} i \lambda T - v_0 A + \frac{2 \kappa \theta}{\sigma^2} D \right)
\]  

(1.110)

with

\[
A = \frac{A_1}{A_2}, \\
A_1 = (\lambda^2 + i \lambda) \sinh \left( \frac{d_H T}{2} \right), \\
A_2 = d_H \cosh \left( \frac{d_H T}{2} \right) + e_T \sinh \left( \frac{d_H T}{2} \right), \\
B = \frac{d_H e^\frac{w_T}{2}}{A_2}, \\
D = \log (B).
\]

(1.111)  (1.112)  (1.113)  (1.114)  (1.115)
Appendix 1.A Heston and Bi-Heston approximation of WSVM

The main advantage of the new representation, apart from the improved stability and lack of discontinuities with respect to alternative formulations, is that it is easily differentiable. In particular, it is possible to compute (semi-)analytically the derivatives of call options model prices \( C^H(\pi_H, K, T) \) with respect to model parameters. Let \( \nabla C^H(\pi_H, K, T) \) be the gradient of \( C^H(\pi_H, K, T) \) with respect to \( \pi_H \), then the following holds

\[
\nabla C^H(\pi_H, K, T) = \frac{1}{\pi} \left( \int_0^\infty \Re \left( \frac{K^{t-x}}{t} \nabla \phi_y^H(\lambda - t, T) \right) d\lambda \right.
\]

\[
- K \int_0^\infty \Re \left( \frac{K^{t-x}}{t} \nabla \phi_y^H(\lambda, T) \right) d\lambda \right) \quad \text{(1.116)}
\]

where \( \nabla \phi_y^H(\lambda, T) = \phi_y^H(\lambda, T)h(\lambda), h(\lambda) = [h_1(\lambda), ..., h_5(\lambda)]^T \) with elements

\[
h_1(\lambda) = -A, \quad \text{(1.117)}
\]

\[
h_2(\lambda) = \frac{\nu_0}{\sigma \lambda} \frac{\partial A}{\partial \rho} + \frac{2\theta}{\sigma^2} D + \frac{2\kappa}{\sigma^2 B} \frac{\partial B}{\partial \kappa} - \frac{\theta \rho \lambda T}{\sigma}, \quad \text{(1.118)}
\]

\[
h_3(\lambda) = \frac{2\kappa}{\sigma^2} D - \frac{\kappa \rho \lambda T}{\sigma}, \quad \text{(1.119)}
\]

\[
h_4(\lambda) = -\nu_0 \frac{\partial A}{\partial \sigma} - \frac{4\kappa \theta}{\sigma^3} D + \frac{2\kappa \theta}{\sigma^2 d_H} \left( \frac{\partial d_H}{\partial \sigma} - \frac{d_H}{A_2} \frac{\partial A_2}{\partial \sigma} \right) + \frac{\kappa \theta \rho \lambda T}{\sigma^2}, \quad \text{(1.120)}
\]

\[
h_5(\lambda) = -\nu_0 \frac{\partial A}{\partial \rho} + \frac{2\kappa \theta}{\sigma^2 d_H} \left( \frac{\partial d_H}{\partial \rho} - \frac{d_H}{A_2} \frac{\partial A_2}{\partial \rho} \right) - \frac{\kappa \theta \lambda T}{\sigma}. \quad \text{(1.121)}
\]

The derivatives appearing in (1.117)-(1.121) can be easily obtained from the definition of involved quantities. We refer to the original paper [30] for their exact formulation.

1.A.2 Jacobian matrix of mapping between WSVM and Heston parameters

Let \( g^{H-W}(\pi_W, T) \) be the mapping presented in Section 1.3.1 between the WSVM and Heston parameters. Without loss of generality we consider \( d = 2 \) and assume matrix \( M \) to be symmetric. For a fixed \( T > 0 \) we have (from Proposition 4 and (1.44)):

\[
v_0 = \text{Tr} [\Sigma_0], \quad \text{(1.122)}
\]

\[
\theta = -\frac{\beta}{2} \text{Tr} [M^{-1} Q^T Q], \quad \text{(1.123)}
\]

\[
\kappa = -\frac{1}{\tau} \log \left( \frac{\mathbb{E} [\mathcal{V}(T)|\Sigma(t)] - \theta}{\mathcal{V}(t) - \theta} \right), \quad \text{(1.124)}
\]

\[
\eta = \sqrt{a \kappa (1 - a) \mathcal{V}(t) + \frac{a}{2} \theta}, \quad \text{(1.125)}
\]
where we simplify the formula for $\theta$ by exploiting the cyclic property of the trace, the symmetry of $M$ and the fact that $\frac{1}{\alpha} M^{-1} e^{\alpha s M}$ is a primitive of $e^{\alpha s M}$.

We now show how to compute the elements of matrix $J_{H - W}(\pi_W)$ defined by (1.59).

Firstly, we recall a useful Lemma about the differentiation of matrix exponentials with respect to scalar parameters:

**Lemma 2** (Wilcox’s formula [93]). Let $A = (\alpha_{i,j})$ be a square matrix of arbitrary order whose elements are functions of (at least) one scalar parameter $p$. It holds that

$$\frac{\partial}{\partial p} \exp(\tau A) = \int_0^\tau \exp((\tau - u)A) A' p \exp(uA) du$$

where $A' = \left( \frac{\partial \alpha_{i,j}}{\partial p} \right)$.

From [91], we know that integrals like (1.127) can be computed via matrix exponentiation: let $C$ be the block matrix $C = \begin{pmatrix} AA' & 0 \\ 0 & A \end{pmatrix}$ and $D = \exp(\tau C) = \begin{pmatrix} D_{1,1} & D_{1,2} \\ 0 & D_{2,2} \end{pmatrix}$, then integral in (1.127) is given by $D_{1,2}$.

Let us now consider the matrices $A_W = \begin{pmatrix} -M Q^\top Q \\ 0 \\ 0 \\ A_W \end{pmatrix}$ and $B_W = \exp(\tau A_W) = \begin{pmatrix} B_{1,1} & B_{1,2} \\ 0 & B_{2,2} \end{pmatrix}$.

From [91], it holds that $\Gamma(\tau) = B_{2,2}^\top \Sigma B_{2,2}$ and $\Theta(\tau) = B_{2,2}^\top B_{1,2}$. For the $q$-th parameter in $\pi_W$ we set $C_W = \begin{pmatrix} A_W & \frac{\partial A_W}{\partial \pi_W,q} \\ 0 & A_W \end{pmatrix}$ and $D_W = \exp(\tau C_W) = \begin{pmatrix} D_{1,1} & D_{1,2} \\ 0 & D_{2,2} \end{pmatrix}$, then from (1.127) and [91], we have that

$$\frac{\partial}{\partial p} \exp(\tau A_W) = D_{1,2}^W = \begin{pmatrix} \frac{\partial B_{1,1}}{\partial \pi_W,q} & \frac{\partial B_{1,2}}{\partial \pi_W,q} \\ 0 & \frac{\partial B_{2,2}}{\partial \pi_W,q} \end{pmatrix}.$$  

Let $e_{i,j} = (1_{k=i,\ell=j})_{1 \leq k,\ell \leq d}$ for $1 \leq i, j \leq d$ form the standard basis of $\mathcal{M}_d(\mathbb{R})$. We can finally write

$$\frac{\partial}{\partial \Sigma_{i,j}} \Gamma(\tau) = B_{2,2}^\top e_{i,j} B_{2,2},$$

$$\frac{\partial}{\partial M_{i,j}} \Gamma(\tau) = \frac{\partial B_{2,2}^\top}{\partial M_{i,j}} \Sigma B_{2,2} + B_{2,2}^\top \Sigma \frac{\partial B_{2,2}}{\partial M_{i,j}},$$

$$\frac{\partial}{\partial M_{i,j}} \Theta(\tau) = \frac{\partial B_{2,2}^\top}{\partial M_{i,j}} B_{1,2} + B_{2,2}^\top \frac{\partial B_{1,2}^\top}{\partial M_{i,j}},$$

$$\frac{\partial}{\partial Q_{i,j}} \Theta(\tau) = B_{2,2}^\top \frac{\partial B_{1,2}^\top}{\partial Q_{i,j}},$$

since $B_{2,2}^\top = \exp(\tau M^\top)$. 

50
Let $\delta_{ij}$ be the Kronecker’s delta, after some algebraic manipulations, the elements of $J^{H-W}(\pi_W)$ can be computed as follows:

$$\frac{\partial v_0}{\partial \Sigma_{i,j}} = \delta_{i,j}, \quad (1.133)$$

$$\frac{\partial \theta}{\partial \beta} = -\frac{1}{2} \text{Tr} \left[ M^{-1} Q^T Q \right], \quad (1.134)$$

$$\frac{\partial \theta}{\partial M_{i,j}} = -\frac{\beta}{2} \text{Tr} \left[ M^{-1} e_{i,j} M^{-1} Q^T Q \right], \quad (1.135)$$

$$\frac{\partial \theta}{\partial Q_{i,j}} = -\frac{\beta}{2} \text{Tr} \left[ M^{-1} \left( e_{j,i} Q + Q^T e_{i,j} \right) \right], \quad (1.136)$$

$$\frac{\partial \kappa}{\partial \beta} = -d_0 \text{Tr} \left[ \Theta(\tau) \right] + d_1 \frac{\partial \Theta}{\partial \beta}, \quad (1.137)$$

$$\frac{\partial \kappa}{\partial \Sigma_{i,j}} = -d_0 \text{Tr} \left[ \frac{\partial \Gamma(\tau)}{\partial \Sigma_{i,j}} \right] - d_2 \delta_{i,j} \frac{\partial \kappa}{\partial \beta}, \quad (1.138)$$

$$\frac{\partial \kappa}{\partial M_{i,j}} = -d_0 \text{Tr} \left[ \frac{\partial \Gamma(\tau)}{\partial M_{i,j}} \right] + d_1 \frac{\partial \Theta}{\partial M_{i,j}} \frac{\partial \kappa}{\partial \beta}, \quad (1.139)$$

$$\frac{\partial \kappa}{\partial Q_{i,j}} = -d_0 \text{Tr} \left[ \frac{\partial \Gamma(\tau)}{\partial Q_{i,j}} \right] - d_2 \delta_{i,j} \frac{\partial \kappa}{\partial \beta} - \text{Var} \left[ V(T) | \Sigma(t) \right] \frac{\partial \eta}{\partial \Sigma_{i,j}} \frac{\partial \kappa}{\partial \beta}, \quad (1.140)$$

$$\frac{\partial \rho}{\partial \Sigma_{i,j}} = \left( \frac{\text{Tr} \left[ R Q e_{i,j} \right]}{\text{Tr} \left[ R Q \Sigma \right]} - \frac{\text{Tr} \left[ \Sigma \right] \text{Tr} \left[ Q^T Q e_{i,j} \right]}{2 \text{Tr} \left[ \Sigma \right] \text{Tr} \left[ Q^T Q \Sigma \right]} \right) \rho, \quad (1.145)$$
Chapter 1. Wishart Processes in Finance

\[
\frac{\partial \rho}{\partial Q_{i,j}} = \frac{\text{Tr} [R e_{i,j} \Sigma] \text{Tr} [Q^\top Q \Sigma]}{\text{Tr} [\Sigma] \text{Tr} [Q^\top Q \Sigma]} - \frac{\text{Tr} [R Q \Sigma] \text{Tr} [\sqrt{Q^\top Q \Sigma}^3]}{\sqrt{\text{Tr} [\Sigma] \sqrt{\text{Tr} [Q^\top Q \Sigma]^3}}} \tag{1.146}
\]

\[
\frac{\partial \rho}{\partial R_{i,j}} = \frac{\text{Tr} [e_{i,j} Q \Sigma] \text{Tr} [R Q \Sigma] \rho,}{\text{Tr} [R Q \Sigma]} \tag{1.147}
\]

with

\[
d_0 = \mathcal{V}(0) - \theta, \tag{1.148}
\]

\[
d_1 = \mathbb{E} [\mathcal{V}(T) | \Sigma] - \mathcal{V}(0), \tag{1.149}
\]

\[
d_2 = \mathbb{E} [\mathcal{V}(T) | \Sigma] - \theta, \tag{1.150}
\]

\[
\bar{a} = \frac{a ((1 - a) \mathcal{V}(t) + \frac{a}{2} \theta)}{\kappa}; \tag{1.151}
\]

where we left unspecified the derivatives of \( \bar{a} \) (that can be easily computed starting from the definition of the terms involved given in Proposition 4). All the missing elements of \( J_{H-W}(\pi_W) \) are equal to zero.

1.A.3 The Bi-Heston model case: new formulation of log-asset characteristic function

We now extend the approach in [30] to the Bi-Heston model [24] with parameters \( \pi_{BH} = [\pi_1, \pi_2]^\top \) and \( \pi_i = [v_{0,i}, \kappa_i, \theta_i, \sigma_i, \rho_i] \) for \( i = 1, 2 \). We simply need to recognize that the Bi-Heston characteristic function corresponds to

\[
\phi_{BH}(\lambda, T) = \exp \left( i \lambda y(0) + \sum_{i=1}^{2} \frac{\kappa_i \theta_i}{\sigma_i^2} (e_{H,i} + d_{H,i}) T - 2 \log \left( \frac{1 - l_{H,i} e^{d_{H,i}^T}}{1 - l_{H,i} e^{d_{H,i}^T}} \right) \right) + \sum_{i=1}^{2} \frac{v_{0,i}}{\sigma_i^2} (e_{H,i} + d_{H,i}) \frac{1 - e^{d_{H,i}^T}}{1 - l_{H,i} e^{d_{H,i}^T}} \tag{1.152}
\]

where the terms in (1.152) are defined in (1.107)-(1.109) provided that we use the corresponding \( i \)-th subset \( \pi_i \). Similarly to the Heston model, then, we can write \( \phi_{y^{BH}} \) as following:

\[
\phi_{y^{BH}}(\lambda, T) = \exp \left( i \lambda y(0) - \sum_{i=1}^{2} \frac{\kappa_i \theta_i \rho_i}{\sigma_i} t \lambda T - v_{0,i} A_i + \frac{2 \kappa_i \theta_i}{\sigma_i^2} D_i \right) \tag{1.153}
\]

where, once more, all the terms involved are defined as in the Heston case for a proper choice of \( \pi_i \). The last step, namely the computation of \( \nabla C^{BH}(\pi_{BH}, K, T) \), follows straightforwardly by noting that each component of the summation in (1.153) does not depend on subset \( \pi_j \) for \( j \neq i \). Following [30], we have

\[
\nabla C^{BH}(\pi_{BH}, K, T) = \frac{1}{\pi} \left( \int_0^{\infty} \Re \left( \frac{K^{-\lambda}}{t \lambda} \nabla \phi_{y^{BH}}(\lambda - t, T) \right) \ d\lambda \right)
\]

52
Appendix 1.A Heston and Bi-Heston approximation of WSVM

\[-K \int_{0}^{\infty} \Re \left( \frac{K^{-i\lambda}}{e^{i\lambda}} \nabla \phi_y^{BH}(\lambda, T) \right) d\lambda \]  \hspace{1cm} (1.154)

where \( \nabla \phi_y^{BH}(\lambda, T) = \phi_y^{BH}(\lambda, T) h(\lambda) \), \( h(\lambda) = [h_1(\lambda), h_2(\lambda)]^T \), with \( h_i(\lambda) = [h_{1,i}(\lambda), ..., h_{5,i}(\lambda)] \) defined in (1.117)-(1.121) for \( i = 1, 2 \).

1.A.4 Jacobian matrix of mapping between WSVM and Bi-Heston parameters

Let \( g_{BH-W}^{W}(\pi_W, T) \) be the mapping presented in presented in Section 1.3.1 between the WSVM and Bi-Heston parameters. Here we show how to compute explicitly the elements of matrix \( J_{BH-W}^{W}(\pi_W) \).

As done for the Heston case, we consider \( d = 2 \) and assume matrix \( M \) to be symmetric.

For a fixed \( T \), Bi-Heston parameters are given in (1.47)-(1.51). Further, let \( Q \) be the orthogonal matrix that diagonalizes \( \Theta(T) \), that is \( Q^\top \Theta(T) Q = E = \text{diag}[\varepsilon_1, \varepsilon_2] \) with \( \varepsilon_1 \geq \varepsilon_2 \) (i.e. with eigenvalues sorted in descending order). By direct computation the following holds for \( i, j, k = 1, 2 \):

\[ \frac{\partial \zeta_i}{\partial \Sigma_{j,k}} = \text{Tr} \left( Q^\top \frac{\partial \Gamma(\tau)}{\partial \Sigma_{j,k}} Q e_{i,i} \right) \]  \hspace{1cm} (1.155)

\[ \frac{\partial \zeta_i}{\partial M_{j,k}} = \text{Tr} \left[ \left( \frac{\partial Q^\top}{\partial M_{j,k}} \Gamma(\tau) Q + Q^\top \frac{\partial \Gamma(\tau)}{\partial M_{j,k}} Q + Q^\top \Gamma(\tau) \frac{\partial Q}{\partial M_{j,k}} \right) e_{i,i} \right] \]  \hspace{1cm} (1.156)

\[ \frac{\partial \zeta_i}{\partial Q_{j,k}} = \text{Tr} \left[ \left( \frac{\partial Q^\top}{\partial Q_{j,k}} \Gamma(\tau) Q + Q^\top \Gamma(\tau) \frac{\partial Q}{\partial Q_{j,k}} \right) e_{i,i} \right] \]  \hspace{1cm} (1.157)

\[ \frac{\partial \nu_{0,i}}{\partial \Sigma_{j,k}} = \text{Tr} \left[ Q^\top e_{j,k} Q e_{i,i} \right] \]  \hspace{1cm} (1.158)

\[ \frac{\partial \nu_{0,i}}{\partial M_{j,k}} = \text{Tr} \left[ \left( \frac{\partial Q^\top}{\partial M_{j,k}} \Sigma_0 Q + Q^\top \Sigma_0 \frac{\partial Q}{\partial M_{j,k}} \right) e_{i,i} \right] \]  \hspace{1cm} (1.159)

\[ \frac{\partial \nu_{0,i}}{\partial Q_{j,k}} = \text{Tr} \left[ \left( \frac{\partial Q^\top}{\partial Q_{j,k}} \Sigma_0 Q + Q^\top \Sigma_0 \frac{\partial Q}{\partial Q_{j,k}} \right) e_{i,i} \right] \]  \hspace{1cm} (1.160)

\[ \frac{\partial \kappa_i}{\partial \Sigma_{j,k}} = \frac{1}{t} \left( \frac{1}{\nu_{0,i}} \frac{\partial \nu_{0,i}}{\partial \Sigma_{j,k}} - \frac{1}{\zeta_i} \frac{\partial \zeta_i}{\partial \Sigma_{j,k}} \right) \]  \hspace{1cm} (1.161)

\[ \frac{\partial \kappa_i}{\partial M_{j,k}} = \frac{1}{t} \left( \frac{1}{\nu_{0,i}} \frac{\partial \nu_{0,i}}{\partial M_{j,k}} - \frac{1}{\zeta_i} \frac{\partial \zeta_i}{\partial M_{j,k}} \right) \]  \hspace{1cm} (1.162)

\[ \frac{\partial \kappa_i}{\partial Q_{j,k}} = \frac{1}{t} \left( \frac{1}{\nu_{0,i}} \frac{\partial \nu_{0,i}}{\partial Q_{j,k}} - \frac{1}{\zeta_i} \frac{\partial \zeta_i}{\partial Q_{j,k}} \right) \]  \hspace{1cm} (1.163)

\[ \frac{\partial \eta_i}{\partial \Sigma_{j,k}} = \frac{\eta_i (1 + t \kappa_i - e^{t \kappa_i})}{2 \kappa_i (1 - e^{t \kappa_i})} \frac{\partial \kappa_i}{\partial \Sigma_{j,k}} \]  \hspace{1cm} (1.164)
Chapter 1. Wishart Processes in Finance

\[
\frac{\partial \theta_i}{\partial \Sigma_{j,k}} = \frac{\theta_i}{\eta \kappa_i} \left( 2\kappa_i \frac{\partial \eta_i}{\partial \Sigma_{j,k}} - \eta_i \frac{\partial \kappa_i}{\partial \Sigma_{j,k}} \right),
\]
\[
\frac{\partial \theta_i}{\partial M_{j,k}} = \frac{\theta_i}{\kappa_i \eta_i} \left( 2\kappa_i \frac{\partial \eta_i}{\partial M_{j,k}} - \eta_i \frac{\partial \kappa_i}{\partial M_{j,k}} \right),
\]
\[
\frac{\partial \theta_i}{\partial Q_{j,k}} = \frac{\theta_i}{\kappa_i \eta_i} \left( 2\kappa_i \frac{\partial \eta_i}{\partial Q_{j,k}} - \eta_i \frac{\partial \kappa_i}{\partial Q_{j,k}} \right),
\]
\[
\frac{\partial \rho_1}{\partial \Sigma_{1,1}} = -\frac{(Q_{1,1} R_{2,1} + Q_{2,1} R_{1,2}) \Sigma_{1,2}}{\sqrt{Q_{1,1}^2 + Q_{2,1}^2 \Sigma_{1,1}^2}},
\]
\[
\frac{\partial \rho_2}{\partial \Sigma_{2,2}} = -\frac{(Q_{1,2} R_{1,1} + Q_{2,2} R_{1,2}) \Sigma_{1,2}}{\sqrt{Q_{1,2}^2 + Q_{2,2}^2 \Sigma_{2,2}^2}},
\]
\[
\frac{\partial \rho_i}{\partial \Sigma_{1,2}} = \frac{\Sigma_{1,i} \partial \rho_i}{\Sigma_{1,2} \partial \Sigma_{1,i}},
\]
\[
\frac{\partial \rho_1}{\partial Q_{1,1}} = \frac{Q_{2,1} (Q_{2,1} (R_{1,1} \Sigma_{1,1} + R_{2,1} \Sigma_{1,2}) - Q_{1,1} (R_{1,2} \Sigma_{1,1} + R_{2,2} \Sigma_{1,2}))}{(Q_{1,1}^2 + Q_{2,1}^2)^{3/2} \Sigma_{1,1}},
\]
\[
\frac{\partial \rho_2}{\partial Q_{1,2}} = \frac{Q_{2,2} (Q_{2,2} (R_{1,1} \Sigma_{1,1} + R_{2,1} \Sigma_{1,2}) - Q_{1,2} (R_{1,2} \Sigma_{1,1} + R_{2,2} \Sigma_{1,2}))}{(Q_{1,2}^2 + Q_{2,2}^2)^{3/2} \Sigma_{2,2}},
\]
\[
\frac{\partial \rho_i}{\partial R_{j,k}} = \frac{Q_{k,i} \Sigma_{i,j}}{\sqrt{Q_{1,i}^2 + Q_{2,i}^2 \Sigma_{i,j}}},
\]

For the sake of simplicity, we left unspecified the derivatives of matrices $E$ and $Q$ with respect to WSVM parameters. These quantities can be easily computed from the
definition of the matrices involved with the help of a computer algebra system like, for example, Mathematica. Even in this case, missing elements of $J^{BH-W}(\pi_W)$ are equal to zero.
Chapter 1. Wishart Processes in Finance

1.B Jacobian matrix of WASC-Heston parameters mapping

Let $g_i^{H-W_A}(\pi_{WA}, T)$ be the mapping presented in Section 1.4.3 between the WASC and Heston model parameters corresponding to the $i$-th basket component. Here we show how to compute explicitly the elements of matrix $J_{i,T}^{H-W_A}(\pi_{WA})$ defined by (1.95).

For the sake of simplicity, and without loss of generality, we consider $d = 2$. By direct computation the following holds for $i, j, k = 1, 2$:

\[
\frac{\partial \nu_{0,i}}{\partial \Sigma_{j,k}} = \mathbb{1}_{i=j=k}, \tag{1.179}
\]

\[
\frac{\partial \kappa_i}{\partial \Sigma_{j,k}} = \frac{1}{t} \left( -\frac{1}{\gamma_i} \frac{\partial \gamma_i}{\partial \Sigma_{j,k}} + \frac{1}{\Sigma_{j,k}} \mathbb{1}_{i=j=k} \right), \tag{1.180}
\]

\[
\frac{\partial \kappa_i}{\partial M_{j,k}} = -\frac{1}{\gamma_i t} \frac{\partial \gamma_i}{\partial M_{j,k}}, \tag{1.181}
\]

\[
\frac{\partial \eta_i}{\partial \Sigma_{j,k}} = \frac{2 \theta_i}{\eta_i} \left( 1 - (1 + \kappa_i t) e^{-\kappa_i t} \right) \frac{\partial \kappa_i}{\partial \Sigma_{j,k}}, \tag{1.182}
\]

\[
\frac{\partial \theta_i}{\partial \Sigma_{j,k}} = \frac{2 \kappa_i}{4 \kappa_i^2} \frac{\partial \eta_i}{\partial \Sigma_{j,k}} \left( 2 \kappa_i \frac{\partial \eta_i}{\partial \Sigma_{j,k}} - \eta_i \frac{\partial \kappa_i}{\partial \Sigma_{j,k}} \right), \tag{1.185}
\]

\[
\frac{\partial \rho_1}{\partial Q_{1,1}} = \frac{\beta \eta_i}{2 \kappa_i \eta_i} \frac{\partial \eta_i}{\partial Q_{1,1}}, \tag{1.188}
\]

\[
\frac{\partial \rho_1}{\partial Q_{1,2}} = \frac{\beta \eta_i}{2 \kappa_i \eta_i} \frac{\partial \eta_i}{\partial Q_{1,2}}, \tag{1.189}
\]

\[
\frac{\partial \rho_2}{\partial Q_{1,1}} = \frac{2 \theta_i}{\eta_i} \left( 1 - (1 + \kappa_i t) e^{-\kappa_i t} \right) \frac{\partial \kappa_i}{\partial Q_{1,1}}, \tag{1.190}
\]

\[
\frac{\partial \rho_2}{\partial Q_{1,2}} = \frac{2 \theta_i}{\eta_i} \left( 1 - (1 + \kappa_i t) e^{-\kappa_i t} \right) \frac{\partial \kappa_i}{\partial Q_{1,2}}. \tag{1.191}
\]
Appendix 1.B Jacobian matrix of WASC-Heston parameters mapping

\[
\frac{\partial \rho_2}{\partial Q_{2,2}} = \frac{Q_{1,2}(Q_{1,2}r_2 - Q_{2,2}r_1)}{(Q_{1,2}^2 + Q_{2,2}^2)^{3/2}}, \quad (1.192)
\]

\[
\frac{\partial \rho_i}{\partial r_j} = \frac{Q_{j,i}}{\sqrt{Q_{1,i}^2 + Q_{2,i}^2}}. \quad (1.193)
\]

The unspecified elements of \(J^{H-W,A}\) are equal to zero.
Chapter 1. Wishart Processes in Finance

1.C Calibration outputs

1.C.1 Calibration results for WSVM

Calibration of WSVM model to market data for the DAX index

<table>
<thead>
<tr>
<th>Parameter</th>
<th>$\beta \geq 0$</th>
<th>$\beta \geq 1$</th>
</tr>
</thead>
<tbody>
<tr>
<td>BH</td>
<td>CF</td>
<td>BH</td>
</tr>
<tr>
<td>$\beta$</td>
<td>0.3612</td>
<td>0.3287</td>
</tr>
<tr>
<td>$\Sigma_{11}$</td>
<td>0.0639</td>
<td>0.0653</td>
</tr>
<tr>
<td>$\Sigma_{12}$</td>
<td>0</td>
<td>0.0105</td>
</tr>
<tr>
<td>$\Sigma_{22}$</td>
<td>0.0237</td>
<td>0.0213</td>
</tr>
<tr>
<td>$M_{11}$</td>
<td>-0.9879</td>
<td>-1.0793</td>
</tr>
<tr>
<td>$M_{12}$</td>
<td>0</td>
<td>-0.8468</td>
</tr>
<tr>
<td>$M_{22}$</td>
<td>-1.3541</td>
<td>-1.4760</td>
</tr>
<tr>
<td>$Q_{11}$</td>
<td>0.4639</td>
<td>0.4060</td>
</tr>
<tr>
<td>$Q_{12}$</td>
<td>0</td>
<td>0.1623</td>
</tr>
<tr>
<td>$Q_{21}$</td>
<td>0</td>
<td>0.4097</td>
</tr>
<tr>
<td>$Q_{22}$</td>
<td>0.4590</td>
<td>0.4763</td>
</tr>
<tr>
<td>$R_{11}$</td>
<td>-0.6618</td>
<td>-0.7280</td>
</tr>
<tr>
<td>$R_{12}$</td>
<td>0</td>
<td>-0.1718</td>
</tr>
<tr>
<td>$R_{21}$</td>
<td>0</td>
<td>0.6232</td>
</tr>
<tr>
<td>$R_{22}$</td>
<td>-0.6272</td>
<td>-0.5645</td>
</tr>
<tr>
<td>Error Price</td>
<td>4.52E-06</td>
<td>2.45E-06</td>
</tr>
<tr>
<td>Error Vol</td>
<td>6.16E-05</td>
<td>2.85E-05</td>
</tr>
<tr>
<td>Time (s)</td>
<td>1.93</td>
<td>24.33</td>
</tr>
</tbody>
</table>

Table 1.4: Calibration on February, 3 2016 with the WSVM over a full set of DAX European call options. As in [32] Error Price stands for the Mean Squared Error (MSE) in price normalized by the forward price and Error Vol denotes the MSE in implied volatility. The part $\beta \geq 0$ shows calibrated parameters for the case with no constraints on parameter $\beta$ while the part $\beta \geq 1$ reports values of parameters obtained imposing the condition of existence and uniqueness of a weak solution to SDE (1.1). Column BH refers to intermediate results obtained via Bi-Heston approximation, while column CF reports final output of calibration routine computed via WSVM characteristic function.
Appendix 1.C Calibration outputs

Calibration of Bi-Heston model to market data for the DAX index

<table>
<thead>
<tr>
<th>Parameter</th>
<th>DAX</th>
</tr>
</thead>
<tbody>
<tr>
<td>$v_{0,1}$</td>
<td>0.0556</td>
</tr>
<tr>
<td>$\kappa_1$</td>
<td>1.6541</td>
</tr>
<tr>
<td>$\theta_1$</td>
<td>0.0534</td>
</tr>
<tr>
<td>$\eta_1$</td>
<td>0.5219</td>
</tr>
<tr>
<td>$\rho_1$</td>
<td>-0.7151</td>
</tr>
<tr>
<td>$v_{0,2}$</td>
<td>0.0312</td>
</tr>
<tr>
<td>$\kappa_2$</td>
<td>1.3922</td>
</tr>
<tr>
<td>$\theta_2$</td>
<td>0.0189</td>
</tr>
<tr>
<td>$\eta_2$</td>
<td>1.5318</td>
</tr>
<tr>
<td>$\rho_2$</td>
<td>-0.6770</td>
</tr>
</tbody>
</table>

| Error Price | 2.83E-06 |
| Error Vol   | 3.56E-05 |
| Time (s)    | 3.58     |

Table 1.5: Calibration on February 3, 2016 with the Bi-Heston over a full set of DAX European call options. As in [32], Error Price stands for the Mean Squared Error (MSE) in price normalized by the forward price and Error Vol denotes the MSE in implied volatility. The calibration of model parameters has been carried out by means of the extension of algorithm in [30] introduced in Section 1.A.3.


Chapter 1. Wishart Processes in Finance

1.C.2 Calibration results for WASC

Bi-dimensional market example

<table>
<thead>
<tr>
<th>Parameter</th>
<th>True Values</th>
<th>Initial Guess</th>
<th>Calibrated Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta$</td>
<td>0.7</td>
<td>1</td>
<td>0.6951</td>
</tr>
<tr>
<td>$\Sigma_{11}$</td>
<td>0.04</td>
<td>0.1</td>
<td>0.0402</td>
</tr>
<tr>
<td>$\Sigma_{12}$</td>
<td>0.03</td>
<td>0</td>
<td>0.0349</td>
</tr>
<tr>
<td>$\Sigma_{22}$</td>
<td>0.04</td>
<td>0.1</td>
<td>0.0399</td>
</tr>
<tr>
<td>$M_{11}$</td>
<td>-0.7</td>
<td>-2</td>
<td>-0.6791</td>
</tr>
<tr>
<td>$M_{12}$</td>
<td>-0.3</td>
<td>-1</td>
<td>-0.2476</td>
</tr>
<tr>
<td>$M_{21}$</td>
<td>-0.05</td>
<td>-1</td>
<td>-0.0836</td>
</tr>
<tr>
<td>$M_{22}$</td>
<td>-1.2</td>
<td>-2</td>
<td>-1.1417</td>
</tr>
<tr>
<td>$Q_{11}$</td>
<td>0.3</td>
<td>0.1</td>
<td>0.2756</td>
</tr>
<tr>
<td>$Q_{12}$</td>
<td>0.3</td>
<td>0</td>
<td>0.3201</td>
</tr>
<tr>
<td>$Q_{21}$</td>
<td>0.2</td>
<td>0</td>
<td>0.2187</td>
</tr>
<tr>
<td>$Q_{22}$</td>
<td>0.3</td>
<td>0.1</td>
<td>0.2976</td>
</tr>
<tr>
<td>$r_1$</td>
<td>-0.6</td>
<td>0</td>
<td>-0.6490</td>
</tr>
<tr>
<td>$r_2$</td>
<td>-0.1</td>
<td>0</td>
<td>-0.0912</td>
</tr>
</tbody>
</table>

Table 1.6: Results for the calibration exercise described in section 1.4.2
Calibration of restricted WASC model to market data for the pair EuroStoxx50-DAX

<table>
<thead>
<tr>
<th>Parameter</th>
<th>$\beta \geq 0$</th>
<th>$\beta \geq 1$</th>
<th>$\beta \geq 3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta$</td>
<td>0.8577</td>
<td>1.0389</td>
<td>3.0110</td>
</tr>
<tr>
<td>$\Sigma_{11}$</td>
<td>0.0697</td>
<td>0.0554</td>
<td>0.0556</td>
</tr>
<tr>
<td>$\Sigma_{12}$</td>
<td>0.0765</td>
<td>0.0654</td>
<td>0.0643</td>
</tr>
<tr>
<td>$\Sigma_{22}$</td>
<td>0.0890</td>
<td>0.0817</td>
<td>0.0788</td>
</tr>
<tr>
<td>$M_{11}$</td>
<td>-1.9763</td>
<td>-0.4626</td>
<td>-0.6740</td>
</tr>
<tr>
<td>$M_{22}$</td>
<td>-1.1605</td>
<td>-0.7311</td>
<td>-0.8590</td>
</tr>
<tr>
<td>$Q_{11}$</td>
<td>0.4422</td>
<td>0.1977</td>
<td>0.1259</td>
</tr>
<tr>
<td>$Q_{12}$</td>
<td>0.1448</td>
<td>0.0893</td>
<td>0.0920</td>
</tr>
<tr>
<td>$Q_{21}$</td>
<td>0.1075</td>
<td>0.0004</td>
<td>0.0440</td>
</tr>
<tr>
<td>$Q_{22}$</td>
<td>0.3843</td>
<td>0.2787</td>
<td>0.1529</td>
</tr>
<tr>
<td>$r_1$</td>
<td>-0.5145</td>
<td>-0.4888</td>
<td>-0.4706</td>
</tr>
<tr>
<td>$r_2$</td>
<td>-0.5247</td>
<td>-0.5262</td>
<td>-0.8490</td>
</tr>
</tbody>
</table>

| Error Price | 2.67E-06 | 3.58E-06 | 7.66E-06 |
| Error Price | 2.97E-06 | 3.16E-06 | 6.11E-06 |
| Error Vol   | 2.62E-04 | 3.25E-04 | 5.20E-04 |
| Error Vol   | 9.88E-05 | 1.53E-04 | 8.33E-04 |
| Time (s)    | 2.81     | 0.88     | 0.92     |

Table 1.7: Calibration on February, 3 2016 with the WASC over a full set of EuroStoxx50 and DAX indices European call options. As in [32] Error Price stands for the Mean Squared Error (MSE) in price normalized by the forward price and Error Vol denotes the MSE in implied volatility. For each column, the first Error Price value refers to the EuroStoxx50 index while the second one to the DAX index. The same applies for the Error Vol values. In all cases we set $\Sigma_{12}$ so that the initial value of cross-asset correlation is equal to 0.9715 (the one year historical correlation computed with daily market data).
Chapter 1. Wishart Processes in Finance

Calibration of Heston model to market data for the pair EuroStoxx50-DAX

<table>
<thead>
<tr>
<th>Parameter</th>
<th>EuroStoxx50</th>
<th>DAX</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\nu_0$</td>
<td>0.0681</td>
<td>0.0842</td>
</tr>
<tr>
<td>$\kappa$</td>
<td>1.5892</td>
<td>1.4568</td>
</tr>
<tr>
<td>$\theta$</td>
<td>0.0450</td>
<td>0.0622</td>
</tr>
<tr>
<td>$\eta$</td>
<td>0.7398</td>
<td>0.6835</td>
</tr>
<tr>
<td>$\rho$</td>
<td>-0.4401</td>
<td>-0.5998</td>
</tr>
<tr>
<td>Error Price</td>
<td>3.04E-06</td>
<td>2.95E-06</td>
</tr>
<tr>
<td>Error Vol</td>
<td>2.20E-04</td>
<td>1.07E-04</td>
</tr>
<tr>
<td>Time (s)</td>
<td>0.15</td>
<td>0.11</td>
</tr>
</tbody>
</table>

Table 1.8: Calibration on February, 3 2016 with the Heston over a full set of EuroStoxx50 and DAX indices European call options. As in [32], Error Price stands for the Mean Squared Error (MSE) in price normalized by the forward price and Error Vol denotes the MSE in implied volatility. The calibration of model parameters has been carried out by means of the algorithm in [30].
Appendix 1.D Numerical results for new simulation schemes

1.D Numerical results for new simulation schemes

1.D.1 WSVM case

We report the prices of European call options obtained with the new simulation schemes proposed for the WSVM. We set the number of simulations to $2 \times 10^5$. Parameters set is given by the rightmost column ($\beta \geq 1$ - Step CF) in Appendix 1.C. Additional data: $f_0 = 100$ and $r = q = 0$. Reference values are computed via COS method. The asterisk means that the corresponding reference value lies outside of the 95% confidence interval.

<table>
<thead>
<tr>
<th>Strike</th>
<th>Reference Value</th>
<th>No. of time steps</th>
<th>MC estimates</th>
<th>Confidence Interval (95%)</th>
<th>Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>$K = 70$</td>
<td>30.6457</td>
<td>5</td>
<td>30.6313</td>
<td>30.5584 - 30.7042</td>
<td>0.05%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>10</td>
<td>30.6771</td>
<td>30.6041 - 30.7502</td>
<td>-0.10%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>50</td>
<td>30.6089</td>
<td>30.5359 - 30.6818</td>
<td>0.12%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>100</td>
<td>30.6945</td>
<td>30.6215 - 30.7676</td>
<td>-0.16%</td>
</tr>
<tr>
<td>$K = 100$</td>
<td>7.1533</td>
<td>5</td>
<td>7.1179</td>
<td>7.0765 - 7.1593</td>
<td>0.50%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>10</td>
<td>7.1571</td>
<td>7.1156 - 7.1987</td>
<td>-0.05%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>50</td>
<td>7.1148</td>
<td>7.0734 - 7.1561</td>
<td>0.54%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>100</td>
<td>7.1679</td>
<td>7.1264 - 7.2094</td>
<td>-0.20%</td>
</tr>
<tr>
<td>$K = 130$</td>
<td>0.1879</td>
<td>5</td>
<td>0.1913</td>
<td>0.1843 - 0.1982</td>
<td>-1.79%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>10</td>
<td>0.1949</td>
<td>0.1879 - 0.2020</td>
<td>-3.73%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>50</td>
<td>0.1875</td>
<td>0.1806 - 0.1943</td>
<td>0.23%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>100</td>
<td>0.1931</td>
<td>0.1860 - 0.2001</td>
<td>-2.73%</td>
</tr>
</tbody>
</table>

Table 1.9: SIA scheme, $T = 0.5$

<table>
<thead>
<tr>
<th>Strike</th>
<th>Reference Value</th>
<th>No. of time steps</th>
<th>MC estimates</th>
<th>Confidence Interval (95%)</th>
<th>Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>$K = 70$</td>
<td>30.6457</td>
<td>5</td>
<td>30.3970*</td>
<td>30.3240 - 30.4701</td>
<td>0.81%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>10</td>
<td>30.4871*</td>
<td>30.4142 - 30.5601</td>
<td>0.52%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>50</td>
<td>30.6208</td>
<td>30.5477 - 30.6939</td>
<td>0.08%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>100</td>
<td>30.6510</td>
<td>30.5778 - 30.7241</td>
<td>-0.02%</td>
</tr>
<tr>
<td>$K = 100$</td>
<td>7.1533</td>
<td>5</td>
<td>6.9670*</td>
<td>6.9248 - 7.0093</td>
<td>2.60%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>10</td>
<td>7.0240*</td>
<td>6.9822 - 7.0658</td>
<td>1.81%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>50</td>
<td>7.1260</td>
<td>7.0842 - 7.1677</td>
<td>0.38%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>100</td>
<td>7.1527</td>
<td>7.1110 - 7.1944</td>
<td>0.01%</td>
</tr>
<tr>
<td>$K = 130$</td>
<td>0.1879</td>
<td>5</td>
<td>0.2252*</td>
<td>0.2175 - 0.2328</td>
<td>-19.83%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>10</td>
<td>0.2054*</td>
<td>0.1982 - 0.2126</td>
<td>-9.32%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>50</td>
<td>0.1995*</td>
<td>0.1923 - 0.2067</td>
<td>-6.16%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>100</td>
<td>0.1933</td>
<td>0.1863 - 0.2003</td>
<td>-2.86%</td>
</tr>
</tbody>
</table>

Table 1.10: GVA scheme, $T = 0.5$
### Table 1.11: SIA scheme, $T = 1$

<table>
<thead>
<tr>
<th>Strike $K$</th>
<th>Reference Value</th>
<th>No. of time steps</th>
<th>MC estimates</th>
<th>Confidence Interval (95%)</th>
<th>Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>$K = 70$</td>
<td>31.7060</td>
<td>5</td>
<td>31.7382</td>
<td>31.6466 - 31.8298</td>
<td>-0.10%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>10</td>
<td>31.7451</td>
<td>31.6535 - 31.8367</td>
<td>-0.12%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>50</td>
<td>31.6892</td>
<td>31.5979 - 31.7805</td>
<td>0.05%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>100</td>
<td>31.7192</td>
<td>31.6274 - 31.8109</td>
<td>-0.04%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>200</td>
<td>31.7295</td>
<td>31.6378 - 31.8211</td>
<td>-0.07%</td>
</tr>
<tr>
<td>$K = 100$</td>
<td>9.5468</td>
<td>5</td>
<td>9.5022</td>
<td>9.4455 - 9.5588</td>
<td>0.47%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>10</td>
<td>9.5195</td>
<td>9.4629 - 9.5761</td>
<td>0.29%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>50</td>
<td>9.4606*</td>
<td>9.4044 - 9.5169</td>
<td>0.90%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>100</td>
<td>9.5252</td>
<td>9.4659 - 9.5792</td>
<td>0.25%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>200</td>
<td>9.5051</td>
<td>9.4485 - 9.5618</td>
<td>0.44%</td>
</tr>
<tr>
<td>$K = 130$</td>
<td>0.8632</td>
<td>5</td>
<td>0.8690</td>
<td>0.8509 - 0.8871</td>
<td>-0.67%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>10</td>
<td>0.8614</td>
<td>0.8434 - 0.8794</td>
<td>0.21%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>50</td>
<td>0.8416*</td>
<td>0.8239 - 0.8593</td>
<td>2.50%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>100</td>
<td>0.8641</td>
<td>0.8461 - 0.8820</td>
<td>-0.10%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>200</td>
<td>0.8711</td>
<td>0.8530 - 0.8892</td>
<td>-0.91%</td>
</tr>
</tbody>
</table>

### Table 1.12: GVA scheme, $T = 1$

<table>
<thead>
<tr>
<th>Strike $K$</th>
<th>Reference Value</th>
<th>No. of time steps</th>
<th>MC estimates</th>
<th>Confidence Interval (95%)</th>
<th>Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>$K = 70$</td>
<td>31.7060</td>
<td>5</td>
<td>31.0803*</td>
<td>30.9883 - 31.1722</td>
<td>1.97%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>10</td>
<td>31.3390*</td>
<td>31.2472 - 31.4308</td>
<td>1.16%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>50</td>
<td>31.5820*</td>
<td>31.4901 - 31.6739</td>
<td>0.39%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>100</td>
<td>31.6393</td>
<td>31.5474 - 31.7311</td>
<td>0.21%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>200</td>
<td>31.7247</td>
<td>31.6329 - 31.8166</td>
<td>-0.06%</td>
</tr>
<tr>
<td>$K = 100$</td>
<td>9.5468</td>
<td>5</td>
<td>9.1327*</td>
<td>9.0752 - 9.1901</td>
<td>4.34%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>10</td>
<td>9.2961*</td>
<td>9.2392 - 9.3530</td>
<td>2.63%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>50</td>
<td>9.4654*</td>
<td>9.4087 - 9.5221</td>
<td>0.85%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>100</td>
<td>9.4879*</td>
<td>9.4312 - 9.5446</td>
<td>0.62%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>200</td>
<td>9.5379</td>
<td>9.4812 - 9.5946</td>
<td>0.09%</td>
</tr>
<tr>
<td>$K = 130$</td>
<td>0.8632</td>
<td>5</td>
<td>0.9409*</td>
<td>0.9219 - 0.9599</td>
<td>-9.00%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>10</td>
<td>0.8870*</td>
<td>0.8687 - 0.9052</td>
<td>-2.75%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>50</td>
<td>0.8683</td>
<td>0.8502 - 0.8863</td>
<td>-0.58%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>100</td>
<td>0.8670</td>
<td>0.8489 - 0.8850</td>
<td>-0.43%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>200</td>
<td>0.8671</td>
<td>0.8493 - 0.8849</td>
<td>-0.45%</td>
</tr>
</tbody>
</table>
## Appendix 1.D Numerical results for new simulation schemes

<table>
<thead>
<tr>
<th>Strike</th>
<th>Reference Value</th>
<th>No. of time steps</th>
<th>MC estimates</th>
<th>Confidence Interval (95%)</th>
<th>Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>$K = 70$</td>
<td>34.8315</td>
<td>10</td>
<td>34.6597*</td>
<td>34.5191 - 34.8003</td>
<td>0.49%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>50</td>
<td>34.7263</td>
<td>34.5861 - 34.8664</td>
<td>0.30%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>100</td>
<td>34.7686</td>
<td>34.6279 - 34.9093</td>
<td>0.18%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>200</td>
<td>34.6135*</td>
<td>34.4732 - 34.7538</td>
<td>0.63%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>300</td>
<td>34.7843</td>
<td>34.6436 - 34.9251</td>
<td>0.14%</td>
</tr>
<tr>
<td>$K = 100$</td>
<td>15.5618</td>
<td>10</td>
<td>15.3539*</td>
<td>15.2524 - 15.4554</td>
<td>1.34%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>50</td>
<td>15.3699*</td>
<td>15.2689 - 15.4708</td>
<td>1.23%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>100</td>
<td>15.4192*</td>
<td>15.3176 - 15.5208</td>
<td>0.92%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>200</td>
<td>15.3212*</td>
<td>15.2202 - 15.4223</td>
<td>1.55%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>300</td>
<td>15.4521*</td>
<td>15.3505 - 15.5537</td>
<td>0.71%</td>
</tr>
<tr>
<td>$K = 130$</td>
<td>5.0151</td>
<td>10</td>
<td>4.8863*</td>
<td>4.8269 - 4.9458</td>
<td>2.57%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>50</td>
<td>4.8624*</td>
<td>4.8036 - 4.9212</td>
<td>3.04%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>100</td>
<td>4.9128*</td>
<td>4.8532 - 4.9724</td>
<td>2.04%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>200</td>
<td>4.8558*</td>
<td>4.7969 - 4.9147</td>
<td>3.18%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>300</td>
<td>4.9294*</td>
<td>4.8700 - 4.9888</td>
<td>1.71%</td>
</tr>
</tbody>
</table>

**Table 1.13: SIA scheme, $T = 3$**

<table>
<thead>
<tr>
<th>Strike</th>
<th>Reference Value</th>
<th>No. of time steps</th>
<th>MC estimates</th>
<th>Confidence Interval (95%)</th>
<th>Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>$K = 70$</td>
<td>34.8315</td>
<td>10</td>
<td>34.4422*</td>
<td>34.3000 - 34.5844</td>
<td>1.12%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>50</td>
<td>34.6277*</td>
<td>34.4858 - 34.7697</td>
<td>0.58%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>100</td>
<td>34.8138</td>
<td>34.6723 - 34.9553</td>
<td>0.05%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>200</td>
<td>34.7483</td>
<td>34.6067 - 34.8899</td>
<td>0.24%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>300</td>
<td>34.7592</td>
<td>34.6179 - 34.9004</td>
<td>0.21%</td>
</tr>
<tr>
<td>$K = 100$</td>
<td>15.5618</td>
<td>10</td>
<td>15.3433*</td>
<td>15.2402 - 15.4463</td>
<td>1.40%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>50</td>
<td>15.4601</td>
<td>15.3575 - 15.5628</td>
<td>0.65%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>100</td>
<td>15.5276</td>
<td>15.4253 - 15.6299</td>
<td>0.22%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>200</td>
<td>15.5131</td>
<td>15.4108 - 15.6153</td>
<td>0.31%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>300</td>
<td>15.5005</td>
<td>15.3987 - 15.6024</td>
<td>0.39%</td>
</tr>
<tr>
<td>$K = 130$</td>
<td>5.0151</td>
<td>10</td>
<td>5.0522</td>
<td>4.9913 - 5.1131</td>
<td>-0.74%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>50</td>
<td>5.0407</td>
<td>4.9803 - 5.1011</td>
<td>-0.51%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>100</td>
<td>5.0232</td>
<td>4.9633 - 5.0831</td>
<td>-0.16%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>200</td>
<td>5.0096</td>
<td>4.9497 - 5.0695</td>
<td>0.11%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>300</td>
<td>4.9801</td>
<td>4.9207 - 5.0395</td>
<td>0.70%</td>
</tr>
</tbody>
</table>

**Table 1.14: GVA scheme, $T = 3$**
### 1. D. 2 WASC case

<table>
<thead>
<tr>
<th>Strike</th>
<th>Reference Value</th>
<th>No. of time steps</th>
<th>MC estimates</th>
<th>Confidence Interval (95%)</th>
<th>Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>$K = 70$</td>
<td>30.8462</td>
<td>10</td>
<td>30.8737</td>
<td>30.8198 - 30.9276</td>
<td>-0.09%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>20</td>
<td>30.8160</td>
<td>30.7625 - 30.8696</td>
<td>0.10%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>50</td>
<td>30.8513</td>
<td>30.7978 - 30.9048</td>
<td>-0.02%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>100</td>
<td>30.8318</td>
<td>30.7783 - 30.8853</td>
<td>0.05%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>200</td>
<td>30.8249</td>
<td>30.7715 - 30.8783</td>
<td>0.07%</td>
</tr>
<tr>
<td>$K = 100$</td>
<td>8.3780</td>
<td>10</td>
<td>8.4584*</td>
<td>8.4256 - 8.4912</td>
<td>-0.96%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>20</td>
<td>8.4073</td>
<td>8.3749 - 8.4398</td>
<td>-0.35%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>50</td>
<td>8.3973</td>
<td>8.3649 - 8.4296</td>
<td>-0.23%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>100</td>
<td>8.3668</td>
<td>8.3344 - 8.3992</td>
<td>0.13%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>200</td>
<td>8.3705</td>
<td>8.3381 - 8.4028</td>
<td>0.09%</td>
</tr>
<tr>
<td>$K = 130$</td>
<td>0.5467</td>
<td>10</td>
<td>0.5869*</td>
<td>0.5786 - 0.5951</td>
<td>-7.34%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>20</td>
<td>0.5696*</td>
<td>0.5616 - 0.5775</td>
<td>-4.17%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>50</td>
<td>0.5573*</td>
<td>0.5495 - 0.5652</td>
<td>-1.94%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>100</td>
<td>0.5470</td>
<td>0.5391 - 0.5549</td>
<td>-0.05%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>200</td>
<td>0.5523</td>
<td>0.5444 - 0.5602</td>
<td>-1.01%</td>
</tr>
</tbody>
</table>

**Table 1.15: Asset 1, TE scheme**

<table>
<thead>
<tr>
<th>Strike</th>
<th>Reference Value</th>
<th>No. of time steps</th>
<th>MC estimates</th>
<th>Confidence Interval (95%)</th>
<th>Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>$K = 70$</td>
<td>30.8462</td>
<td>10</td>
<td>30.3028*</td>
<td>30.2511 - 30.3545</td>
<td>1.76%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>20</td>
<td>30.5859*</td>
<td>30.5333 - 30.6384</td>
<td>0.84%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>50</td>
<td>30.7412*</td>
<td>30.6881 - 30.7943</td>
<td>0.34%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>100</td>
<td>30.8328</td>
<td>30.7796 - 30.8860</td>
<td>0.04%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>200</td>
<td>30.8521</td>
<td>30.7987 - 30.9055</td>
<td>-0.02%</td>
</tr>
<tr>
<td>$K = 100$</td>
<td>8.3780</td>
<td>10</td>
<td>7.8458*</td>
<td>7.8151 - 7.8766</td>
<td>6.35%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>20</td>
<td>8.1059*</td>
<td>8.0744 - 8.1375</td>
<td>3.25%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>50</td>
<td>8.2812*</td>
<td>8.2491 - 8.3132</td>
<td>1.16%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>100</td>
<td>8.3399*</td>
<td>8.3078 - 8.3720</td>
<td>0.45%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>200</td>
<td>8.3799</td>
<td>8.3476 - 8.4122</td>
<td>-0.02%</td>
</tr>
<tr>
<td>$K = 130$</td>
<td>0.5467</td>
<td>10</td>
<td>0.4424*</td>
<td>0.4355 - 0.4494</td>
<td>19.08%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>20</td>
<td>0.4895*</td>
<td>0.4822 - 0.4968</td>
<td>10.47%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>50</td>
<td>0.5240*</td>
<td>0.5164 - 0.5316</td>
<td>4.16%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>100</td>
<td>0.5368*</td>
<td>0.5291 - 0.5445</td>
<td>1.82%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>200</td>
<td>0.5454</td>
<td>0.5375 - 0.5532</td>
<td>0.25%</td>
</tr>
</tbody>
</table>

**Table 1.16: Asset 1, GVA scheme**
### Appendix 1.6 Numerical results for new simulation schemes

<table>
<thead>
<tr>
<th>Strike Value</th>
<th>No. of time steps</th>
<th>MC estimates</th>
<th>Confidence Interval (95%)</th>
<th>Error</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>K</em> = 70</td>
<td>10</td>
<td>31.5774</td>
<td>31.5190 - 31.6359</td>
<td>-0.12%</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>31.4685*</td>
<td>31.4104 - 31.5266</td>
<td>0.23%</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>31.5378</td>
<td>31.4798 - 31.5958</td>
<td>0.01%</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>31.5116</td>
<td>31.4537 - 31.5695</td>
<td>0.09%</td>
</tr>
<tr>
<td></td>
<td>200</td>
<td>31.5149</td>
<td>31.457 - 31.5728</td>
<td>0.08%</td>
</tr>
<tr>
<td><em>K</em> = 100</td>
<td>10</td>
<td>9.7750*</td>
<td>9.7402 - 9.8097</td>
<td>-1.10%</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>9.6683</td>
<td>9.6339 - 9.7027</td>
<td>0.01%</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>9.6837</td>
<td>9.6493 - 9.7181</td>
<td>-0.15%</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>9.6558</td>
<td>9.6215 - 9.6901</td>
<td>0.14%</td>
</tr>
<tr>
<td></td>
<td>200</td>
<td>9.6603</td>
<td>9.6260 - 9.6946</td>
<td>0.09%</td>
</tr>
<tr>
<td><em>K</em> = 130</td>
<td>10</td>
<td>0.6643*</td>
<td>0.6570 - 0.6717</td>
<td>-9.41%</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>0.6234*</td>
<td>0.6164 - 0.6303</td>
<td>-2.66%</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>0.6173*</td>
<td>0.6105 - 0.6242</td>
<td>-1.67%</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>0.6117</td>
<td>0.6049 - 0.6185</td>
<td>-0.74%</td>
</tr>
<tr>
<td></td>
<td>200</td>
<td>0.6091</td>
<td>0.6023 - 0.6158</td>
<td>-0.31%</td>
</tr>
</tbody>
</table>

Table 1.17: *Asset 2, TE scheme*

<table>
<thead>
<tr>
<th>Strike Value</th>
<th>No. of time steps</th>
<th>MC estimates</th>
<th>Confidence Interval (95%)</th>
<th>Error</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>K</em> = 70</td>
<td>10</td>
<td>30.7237*</td>
<td>30.6679 - 30.7794</td>
<td>2.59%</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>31.1076*</td>
<td>31.0509 - 31.1644</td>
<td>1.37%</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>31.3530*</td>
<td>31.2956 - 31.4104</td>
<td>0.59%</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>31.5186</td>
<td>31.461 - 31.5762</td>
<td>0.07%</td>
</tr>
<tr>
<td></td>
<td>200</td>
<td>31.5372</td>
<td>31.4794 - 31.5950</td>
<td>0.01%</td>
</tr>
<tr>
<td><em>K</em> = 100</td>
<td>10</td>
<td>8.8934*</td>
<td>8.8612 - 8.9256</td>
<td>8.02%</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>9.2535*</td>
<td>9.2203 - 9.2868</td>
<td>4.30%</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>9.4986*</td>
<td>9.4648 - 9.5325</td>
<td>1.76%</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>9.6168*</td>
<td>9.5827 - 9.6508</td>
<td>0.54%</td>
</tr>
<tr>
<td></td>
<td>200</td>
<td>9.6496</td>
<td>9.6154 - 9.6838</td>
<td>0.20%</td>
</tr>
<tr>
<td><em>K</em> = 130</td>
<td>10</td>
<td>0.4325*</td>
<td>0.4270 - 0.4381</td>
<td>28.76%</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>0.5132*</td>
<td>0.5071 - 0.5193</td>
<td>15.48%</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>0.5688*</td>
<td>0.5623 - 0.5753</td>
<td>6.32%</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>0.5863*</td>
<td>0.5797 - 0.5929</td>
<td>3.44%</td>
</tr>
<tr>
<td></td>
<td>200</td>
<td>0.6018</td>
<td>0.5951 - 0.6085</td>
<td>0.89%</td>
</tr>
</tbody>
</table>

Table 1.18: *Asset 2, GVA scheme*
## Table 1.19: Best-Of Put options, TE scheme

<table>
<thead>
<tr>
<th>Strike</th>
<th>Reference Value</th>
<th>No. of time steps</th>
<th>MC estimates</th>
<th>Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>$K = 70$</td>
<td>0.7455</td>
<td>10</td>
<td>0.7382</td>
<td>0.98%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>20</td>
<td>0.7518</td>
<td>-0.84%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>50</td>
<td>0.7358</td>
<td>1.30%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>100</td>
<td>0.7351</td>
<td>1.40%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>200</td>
<td>0.7383</td>
<td>0.97%</td>
</tr>
<tr>
<td>$K = 100$</td>
<td>7.0387</td>
<td>10</td>
<td>7.0678</td>
<td>-0.41%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>20</td>
<td>7.0894</td>
<td>-0.72%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>50</td>
<td>7.0292</td>
<td>0.14%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>100</td>
<td>7.0203</td>
<td>0.26%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>200</td>
<td>7.0376</td>
<td>0.02%</td>
</tr>
<tr>
<td>$K = 130$</td>
<td>26.4237</td>
<td>10</td>
<td>26.3951</td>
<td>0.11%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>20</td>
<td>26.4880</td>
<td>-0.24%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>50</td>
<td>26.4200</td>
<td>0.01%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>100</td>
<td>26.4339</td>
<td>-0.04%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>200</td>
<td>26.4471</td>
<td>-0.09%</td>
</tr>
</tbody>
</table>

## Table 1.20: Best-Of Put options, GVA scheme

<table>
<thead>
<tr>
<th>Strike</th>
<th>Reference Value</th>
<th>No. of time steps</th>
<th>MC estimates</th>
<th>Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>$K = 70$</td>
<td>0.7455</td>
<td>10</td>
<td>0.6712</td>
<td>9.97%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>20</td>
<td>0.6886</td>
<td>7.64%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>50</td>
<td>0.7178</td>
<td>3.72%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>100</td>
<td>0.7117</td>
<td>4.53%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>200</td>
<td>0.7198</td>
<td>3.44%</td>
</tr>
<tr>
<td>$K = 100$</td>
<td>7.0387</td>
<td>10</td>
<td>6.9902</td>
<td>0.69%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>20</td>
<td>6.9632</td>
<td>1.07%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>50</td>
<td>6.9998</td>
<td>0.55%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>100</td>
<td>6.9536</td>
<td>1.21%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>200</td>
<td>6.9755</td>
<td>0.90%</td>
</tr>
<tr>
<td>$K = 130$</td>
<td>26.4237</td>
<td>10</td>
<td>27.1112</td>
<td>-2.60%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>20</td>
<td>26.7730</td>
<td>-1.32%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>50</td>
<td>26.6110</td>
<td>-0.71%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>100</td>
<td>26.4748</td>
<td>-0.19%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>200</td>
<td>26.4704</td>
<td>-0.18%</td>
</tr>
</tbody>
</table>
A New Class of Multidimensional Wishart-based hybrid models

2.1 Introduction

Local Volatility (LV) and Stochastic Volatility (SV) are two of the most widely used approaches to overcome some well-known drawbacks of Black-Scholes model. LV, for example, has become a market standard for derivatives pricing and risk management within several asset classes, e.g. equity and FX. In early 1990s milestone papers [39] and [42] open the way to the concept of LV: specifying a unique time and state dependent diffusion process it is possible to consistently reproduce any given European options surface observed on the market. Although the result is quite elegant, it requires the existence of a complete surface of quoted options. However, in practice we only observe prices for a discrete set of instruments with different combinations of moneyness and time to maturity. We then need to rely on techniques able to "complete" the market surface consistently with the absence of arbitrage opportunities (roughly speaking this would mean that the corresponding local variance surface remains non-negative). Some recent papers go in this direction: in [48], easy to check conditions are formulated to eliminate static arbitrages in the famous Stochastic Volatility Inspired (SVI) parametrization [46], while a finite difference method is used in [5] to get a complete surface of arbitrage consistent option values from sparse market prices. Despite the
ability in matching the spot implied volatilities, the LV approach has severe limitations in pricing exotic options. The inherent deterministic nature of volatility does not allow to satisfactorily describe the underlying asset volatility dynamics. A well known consequence is that LV models tend to generate forward implied volatility surfaces much flatter than the market ones (see [47] for an overview on the topic).

The drawbacks of LV are partially overcome by the use of SV models where an additional stochastic process is introduced to describe the dynamics of volatility. A massive research activity has been produced in this field revealing the wide diffusion of such approach both within practitioners and academics. Notwithstanding the parametrization adopted, the more realistic description of volatility in SV models compared to the LV case comes at a price: (diffusion) processes with embedded SV usually are not able to match market implied volatilities in deep far from the money regions for short dated options. The first attempt to improve the ability of SV models to match marginal distribution of stock prices has been the use of jumps as in [10]. Another solution could originate from the observation that LV and SV models are complementary to some extent: this is exactly the idea behind the so-called Stochastic Local Volatility (SLV) or hybrid approach. Combining both LV and SV models could lead to an overall improvement of pricing accuracy since the resulting hybrid model would benefit from a satisfactory description of volatility dynamics (SV) and the ability to reproduce market plain vanilla option prices (LV). A fast growing interest in SLV models has recently led to a widespread research activity aimed at formulating valid approaches. We refer to [65] for a comprehensive overview of the topic. An important feature of SLV models is that it is possible to separately calibrate the LV and SV components and combine them in a non linear and non parametric way by means of the so-called leverage function. The peculiar and most delicate step to perform when dealing with SLV is the calibration of leverage function itself since we need to account for the dependency between LV and SV. Several techniques have been devised to tackle the problem: in [43] and [83] algorithms are presented to numerically solve the corresponding non linear integral Fokker-Planck PDE. This method is well suited for models with embedded 1-factor SV component while for multifactor ones it would suffer from the curse of dimensionality. Recently some simulation based techniques have been proposed: Markovian projection [62], particle method [60] and a non parametric numerical scheme [90]. These last two methods allow to deal easily with multidimensional processes since they basically require only to be able to sample trajectories for the asset price.

A common SLV specification in the single-asset framework is the use of non parametric local volatility combined with a Heston-like stochastic volatility process as in [89] and [90]. However, in the light of the poor ability of 1-factor SV models in dealing with skew dynamics, as stated in Chapter 1, it seems appropriate to introduce a richer structure for the SV part. To reach such a goal, we present a new hybrid model, the Wishart Stochastic Local Volatility model (WSLVM).
Moreover, the pricing of derivatives with a multi-asset risk exposure requires a satisfactory modelling of the correlation structure among the involved assets, consistently with an accurate reproduction of individual plain vanilla market evidences. This leads to the introduction of multi-asset SLV models as done in [38], where a simplified multi-asset Heston framework is extended to accommodate for hybrid asset covariations. In order to incorporate a more flexible dependence structure among state variables, we introduce the Wishart Stochastic Local Covariance model (WSLCM) as hybrid generalization of the pure SV model in [33]. As far as we know, this new class of models represents one of the first attempts, and probably the most comprehensive one, to provide a multidimensional variance dynamics within the SLV pricing paradigm.

Besides the extension of SLV literature to the case of matrix-based stochastic volatility parametrizations, we aim at improving the pricing performance of pure SV Wishart models. As shown in Chapter 1, indeed, we must introduce stringent parameters restrictions to satisfy existence and uniqueness conditions for the solution of Wishart SDE. This is needed, for example, when we want to simulate the variance process in order to price more exotic derivatives. However, such conditions are not usually met when market calibration is performed. Consequently, the constrained parameters set obtained enforcing such conditions is not able to accurately reproduce the market implied volatility surface. In our new setting, the additional LV component acts as a compensator meant to reduce the gap with the market still preserving the well definiteness of Wishart processes.

The rest of the Chapter is organized as follows: in Section 2.2 we briefly describe the general SLV framework. In Section 2.3 we present the WSLVM and report a realistic implementation of the model using the market dataset of DAX option prices already presented in Chapter 1. In Section 2.4 we illustrate the main properties of WSLCM and show an application on market data for the pair of indices EuroStoxx50-DAX. Finally, in Section 2.5 we state some concluding remarks.

### 2.2 The general SLV framework

A SLV model specifies the following risk-neutral dynamics for the $T$-forward price $f(t)$ of an equity asset

$$df(t) = f(t) \sigma(t, f(t)) \psi(V(t)) dB(t),$$  \hspace{1cm} (2.1)

where $V(t)$ is a (possibly multifactor) stochastic process that affects the dynamics of $f(t)$ through the functional form $\psi(\cdot)$ and $B(t)$ is a Brownian motion of suitable dimension that can be correlated with the source(s) of randomness in $V(t)$. As an example if $V(t)$ is described in terms of 1-factor CIR process and $\psi(v) = \sqrt{v}$, then we get the popular Heston Stochastic Local Volatility model [89], [90].

In this paper we consider, without loss of generality, risk-free rate and dividend yield
equal to zero. We refer the reader to [59] for the extension of SLV models to the case
of stochastic interest rates and/or discrete dividends.
Here $\sigma(t, f(t))$ is the leverage function that can be interpreted as a (deterministic) local
adjustment factor introduced to match the underlying asset terminal distribution ob-
served on the market. The definition of $\sigma(t, f(t))$ is the key aspect of SLV modelling
as well as the most challenging task in the implementation of any model of this kind.
Before stating the fundamental result about SLV modelling, we recall an important the-
orem in stochastic calculus that represents the main tool to develop the formal ground
of the framework.

**Theorem 2 (Mimicking process).** Let $\xi(t)$ be a stochastic process satisfying the SDE
\[ d\xi(t) = \alpha_0(\omega, t)dt + \alpha_1(\omega, t)dW(t) \]
where $W(t)$ is a (possibly multidimensional) Brownian motion and $\alpha_0$, $\alpha_1$ bounded,
non anticipative processes such that $\alpha_1\alpha_1^\top$ is uniformly positive semi definite. Then
there exists a SDE
\[ dx(t) = a_0(t, x(t))dt + a_1(t, x(t))dW(t) \]
with non random coefficients that admits a weak solution $x(t)$ having the same one-
dimensional distribution as $\xi(t)$ $\forall$ $t$. The coefficients $a_0$ and $a_1$ have the following
representation
\[ a_0(t, x) = \mathbb{E}[\alpha_0(t)|\xi(t) = x], \]
\[ a_1(t, x) = \sqrt{\mathbb{E}\left[\alpha_1(t)\alpha_1^\top(t)|\xi(t) = x\right]} \]

**Proof.** See [61].

Thanks to this result, we are supplied with an explicit strategy to construct stochastic
processes driven by deterministic coefficients that “mimic“ some other (more involved)
processes with stochastic coefficients. We now state a general result, firstly shown
in [83], that links the leverage function to both the LV and SV components.

**Proposition 8 (Leverage function).** Given a SLV model defined by (2.1) the market
observed implied volatility surface is matched if and only if the leverage function is of
the form$^[]$
\[ \sigma(t, f) = \frac{\sigma_{LV}(t, f)}{\sqrt{\mathbb{E}^q[\psi^2(V(t))]|f(t) = f]}, \quad (2.2) \]
where $\sigma_{LV}(t, f)$ is the local volatility calibrated on the same market prices.

$^[]$The expectation in (2.2) is computed under the risk-neutral measure since we are interested in derivatives pricing.
2.2. The general SLV framework

Proof. In its basic intuition, (2.2) is a straightforward application of Theorem 2: let us set

\[ \sigma_{LV}(t, f) = \sqrt{\mathbb{E}^{Q}[\sigma^2(t, f)\psi^2(V(t))|f(t) = f]}, \]

(2.3)
such that the solution of LV process \( df(t) = \sigma_{LV}(t, f)f(t)dB(t) \) and (2.1) share the same marginal distribution for \( f(t) \). The matching with the market follows directly from the definition of local volatility process as given in [42]. We refer the interested reader to the original paper [83] for a complete proof.

Proposition 8 provides an interesting interpretation for the leverage function as the ratio between the local volatility and the expected value of the stochastic volatility component. Intuitively this suggests how a SLV model works: if the (expected) level of stochastic volatility is significantly different to the local volatility, then the leverage function acts as compensator (on average).

Clearly the specification (2.1) includes the "pure" LV or SV models as particular cases: with \( \sigma(t, f) = 1 \) we get a standard SV model while assuming no randomness in \( V(t) \) we end up with the LV dynamics.

As pointed out in [59], substituting (2.2) into (2.1) we get a McKean SDE where the diffusion coefficient depends not only on state variables but also on their joint probability distribution. Keeping on following [59], issues about existence and uniqueness of SLV SDE (2.1) have been raised since we have no certainty at all that any market implied volatility surface can be matched given a set of parameter for the SV part.

Numerical investigations show that the problem becomes quite relevant when large values of volatility of volatility are needed to obtain very steep forward skew in standard (1-factor) SV models. The same observation is found in [87] where jumps are added to the SLV dynamics.

2.2.1 Calibration of a SLV model

In order to be able to implement a given SLV model, we firstly need to calibrate its 3 main components: Local Volatility surface, Stochastic Volatility parameters and leverage function (2.2). Remarkably, it is possible to decompose the overall problem into 3 individual tasks to be tackled separately. In particular, the calibration of leverage function can be carried out via ad hoc algorithms after having determined the LV and SV components. In the following we illustrate the calibration of LV surface and leverage function. The estimate of SV parameters is, indeed, highly influenced by modelling assumptions and requires a customized approach. In our case, we refer to Chapter 1 for an extensive treatment of the methodologies applied to the calibration of Wishart based parameters.
Chapter 2. A New Class of Multidimensional Wishart-based hybrid models

Calibration of LV surface

The LV surface can be estimated following one of the (numerous) approaches proposed in literature. Despite the lack of evidences in favor of a particular calibration scheme, we remark that the chosen approach should not generally affect the outputs of SLV modelling, provided that the method produces realistic results. We choose to consider a non-parametric specification of LV surface and calibrate it using the algorithm in [1].

The scheme, which is basically a slight modification of [5], has been proven to give accurate results in repricing calibration instruments with both PDE and Monte Carlo methods. Let $0 = T_0 < T_1 < T_N$ be a set of option expiry dates and consider the Dupire forward equation [42]

$$\frac{\partial C(K, T)}{\partial T} = \frac{1}{2} \sigma_{LV}(T, K)^2 K^2 \frac{\partial^2 C}{\partial K^2}(K, T)$$

(2.4)

where we set, for the sake of simplicity, interest rates and dividends to zero. In [5], the authors propose to calibrate (2.4) between two consecutive dates with the following one-step implicit finite difference scheme

$$\left[ 1 - \frac{1}{2} (T_{i+1} - T_i) \tilde{\sigma}_i(K)^2 K^2 \frac{\partial^2}{\partial K^2} \right] C(K, T_{i+1}) = C(K, T_i),$$

(2.5)

$$C(K, 0) = [S(0) - K]^+, i = 0, 1, ..., N$$

(2.6)

where $\tilde{\sigma}_i(K)$ are the volatility proxies to be calibrated by minimizing the difference between the market prices and those obtained by (2.5). Further, the function $\tilde{\sigma}(T, K)$ is considered to be piecewise constant in time and strike. The modification proposed in [1] considers the same optimization problem but introduces a finer time grid between market expirations in order to produce more accurate results. The corresponding equation becomes

$$\left[ 1 - \frac{1}{2} \frac{(T_{i+1} - T_i)}{N_T} \tilde{\sigma}_i(K)^2 K^2 \frac{\partial^2}{\partial K^2} \right] C\left(K, T_i + \frac{(T_{i+1} - T_i)}{N_T}\right) = C(K, T_i),$$

(2.7)

$$C(K, 0) = [S(0) - K]^+, i = 0, 1, ..., N$$

(2.8)

with $N_T$ the number of points in the temporal mesh between any two consecutive dates.

An additional advantage of the method (in comparison with the original scheme in [5]) is that the calibrated volatility proxies represent in this case the calibration output with no need of further processing. In Figure 2.1 we report the calibrated LV surfaces for DAX and EuroStoxx50 indices. In both cases, results are obtained in less than 2 seconds.

Calibration of leverage function

Given a LV surface and a set of SV parameters (ideally) calibrated on the same basket of market options, the last step in the implementation of a SLV model is the calibration of (2.2). The task, however, is not trivial since the computation of the conditional
2.2. The general SLV framework

Figure 2.1: Local volatility surfaces calibrated to market data on February 3, 2016. DAX index (left) and EuroStoxx50 index (right).

expectation therein would require to know the joint distribution of \( f(t) \) and \( V(t) \). This is clearly not the case due to the presence of the non-parametric LV component that affects the dynamics of \( f(t) \) itself.

In 1-factor SLV models PDE methods are proposed to solve the corresponding Fokker-Planck equation. However in multifactor SLV models these methods are totally unfeasible since we would deal with high-dimensional PDEs.

Simulation methods appear to be more appropriate and flexible enough to be applied within multifactor SLV frameworks. Among the proposed approaches, the non parametric method presented in [90] is the most intuitive and easiest to implement. Since we extensively use this technique for our numerical tests, a brief description is given. We refer the interested reader to the original paper for a detailed illustration and a complete error analysis.

Let \( \hat{f} \) be the discrete time approximation to the SLV equation (2.1). A naive Euler scheme would take the form

\[
\hat{f}_j(t_{m+1}) = \hat{f}_j(t_m) \sigma(t_m, \hat{f}_j(t_m)) \psi(\hat{V}_j(t_m)) \sqrt{\Delta} B, \quad \hat{f}_j(0) = f_0 > 0, \tag{2.9}
\]

for the generic \( j \)-th simulation, \( j = 1, \ldots, N \), at a given time step \( m + 1 \) with \( m = 0, \ldots, M_T \). Here \( N \) is the total number of simulated paths and \( M_T \) the number of time steps. We also set a uniform time grid \( \Delta = T/M_T \) where \( T \) is the terminal date. The (multidimensional) standard normal variable \( B \) is assumed to be enriched with the correlation structure assumed by the chosen SV model (at this stage, for sake of generality, we leave it unspecified). The stochastic variance process \( V \) is simulated with an appropriate scheme as well.
Substituting (2.2) in (2.9) we get

$$\hat{f}_j(t_{m+1}) = \hat{f}_j(t_m) \frac{\sigma_{LV}(t_m, \hat{f}_j(t_m))}{\sqrt{\mathbb{E}^Q \left[ \psi^2(V(t_m)) \mid f(t_m) = \hat{f}_j(t_m) \right]}} \psi(\hat{V}_j(t_m)) \sqrt{\Delta B}. \quad (2.10)$$

The idea in [90] is to group the $N$ realizations of pair $(\hat{f}_j(t_m), \hat{V}_j(t_m))$ at each time step $m$ into $\ell$ mutually exclusive subsets based on sorted values of $\hat{f}_j(t_m)$. The approximation of conditional expectation in (2.10) is then given by

$$\mathbb{E}^Q \left[ \psi^2(V(t_m)) \mid f(t_m) = \hat{f}_j(t_m) \right] \approx \mathbb{E}^Q \left[ \psi^2(V(t_m)) \mid f(t_m) \in (b_k, b_{k+1}] \right] \quad (2.11)$$

for a proper choice of bins $(b_1, b_2], (b_2, b_3], ..., (b_{\ell}, b_{\ell+1})$ with $b_1 \geq 0$ and $b_{\ell+1} < \infty$. Two different bins specifications are proposed in [90]: equidistant and equally weighted (i.e., into each subset there is approximately the same number of realizations) bins. The latter has been found to give better results in terms of approximation accuracy in particular and it is the one used in our numerical tests. This means that for each time step we choose the $\ell + 1$ bins according to

$$b_1(t_m) = \bar{f}_1(t_m), \ b_{\ell+1}(t_m) = \bar{f}_N(t_m), \ b_k(t_m) = \bar{f}_{(k-1)N/\ell}(t_m), \ k = 2, ..., \ell \quad (2.12)$$

where $\{\hat{f}_j(t_m)\}_{j=1}^N$ indicates the sequence of $\hat{f}_j(t_m)$ sorted in ascending order.

### 2.3 The Wishart Stochastic Local Volatility model

#### 2.3.1 Model dynamics

In order to increase the ability of WSVM in describing the marginal probability distribution of asset price, we introduce the Wishart Stochastic Local Volatility model given by

$$df(t) = f(t) \sigma(t, f(t)) \text{Tr} \left[ \sqrt{\Sigma(t)} dB(t) \right], \quad (2.13)$$

where $\Sigma(t)$ is a Wishart process described by (1.1). Equivalently we can write

$$df(t) = f(t) \sigma(t, f(t)) \sqrt{V(t)} db(t), \quad (2.14)$$

with dynamics of $V(t)$ and stochastic correlation given respectively by (1.38) and (1.35). The latter specification leads to an immediate interpretation of (2.14) as a SLV model in light of the general framework defined by (2.1). That is we get the WSLVM dynamics as soon as we set $V(t) = \sqrt{V(t)}$ and $\psi(v) = \sqrt{v}$ in (2.1).

Consistently with Proposition 8, we define the WSLVM leverage function as

$$\sigma(t, f) = \frac{\sigma_{LV}(t, f)}{\sqrt{\mathbb{E}^Q [\text{Tr} [V(t)] | f(t) = f]}}, \quad (2.15)$$
such that, by Theorem 2, the model matches the market implied probability distribution of asset prices. The resulting model induces a hybrid instantaneous variance for the log-asset \( y(t) = \log(f(t)) \) given by

\[
d [y(t), y(t)] = \sigma^2(t, e^{y(t)}) V(t) dt
\]

with the remarkable property to embed a multi-factor specification of the stochastic volatility component. Equation (2.13) represents, indeed, a generalization of WSVM meant to combine the flexibility granted by the underlying Wishart process with an accurate pricing of plain vanilla options. Even if our modelling choice can be seen questionable since we lose the analytical tractability of WSVM we have numerous arguments in favour. Firstly, as in any SLV model, the calibration of LV and SV is performed separately then we can still rely of fast calibration routines for the SV parameters. Secondly Monte Carlo pricing is unavoidable when we deal with exotic options even in affine models. Finally, as shown in the next Section, we get an evident improvement in the pricing performance of European options that could not be achieved otherwise. We think indeed that moving towards a hybrid framework is the only viable way to improve the accuracy of WSVM given the restrictions on the parameter \( \beta \).

The usual fix of adding jumps presents several drawbacks in this case. The additional increase in the number of parameters comes with no guarantee that the lost of accuracy imposed by the restriction \( \beta \geq d - 1 \) would be generally overcome. Analogously to the 1-factor SV case, apart from anything else, jumps pose non-trivial problems in a risk management perspective. As illustrated in [32] adding jumps on the stock and/or on the volatility could lead to a loss of parameters sensitivity with respect to the skew since it is into some extent transferred to the parameters that drive the discontinuities.

2.3.2 Numerical Results

Pricing routine within the WSLVM follows the same steps as in any other SLV model: 1) calibration of LV and SV components, 2) calibration of leverage function and 3) pricing. In Chapter 1 we have already performed the calibration of Wishart-like SV component over a set of market prices of call options written on DAX index. In particular, we consider here the parameters obtained enforcing the condition \( \beta \geq 1 \) (rightmost column of Table 1.4) in order to deal with a well-defined variance process. Exploiting the simulation based approach in [90] for the calibration of leverage function, we can address steps 2) and 3) in a single run enhancing the computational efficiency of the procedure. In order to discretize (2.14) we can apply one of the new WSVM sampling schemes provided in Section 1.3.3. The use of an efficient scheme, indeed, is greatly recommended in this case: since in the SLV framework we deal with the simulation of a process with an embedded local volatility component, which is known to require a fine time discretization to get precise results, we need to rely on a fast and accurate sampling procedure. We simply need to introduce some slight
modifications to take into account the presence of the additional local adjustment term. Given the evidences of the numerical analysis carried out in the previous Chapter, we implement the GVA scheme that for the WSLVM reads as

\[
\hat{y}(t + \Delta) = \hat{y}(t) - \frac{1}{2} \sigma^2(t, e^{\tilde{y}(t)}) \hat{V}(t) \Delta \\
+ \sigma(t, e^{\tilde{y}(t)}) \left( \hat{\rho}_W(t) \sqrt{\hat{V}(t) \tilde{w}} + \sqrt{(1 - \hat{\rho}_W^2(t)) \hat{V}(t) \tilde{z}} \right).
\]

Further, we set the number of bins required to compute(2.11) to 50 and we choose them accordingly to (2.12) at any time step. The Monte Carlo simulations are performed sampling \(10^5\) trajectories with daily time steps. With this settings we are able to solve the calibration problem for a time horizon of 3 years in less than 28 seconds.\(^2\) Storing the calibrated values of the leverage function as well as the chosen bins we can reuse them for further applications.

We are finally ready to test the ability of WSLVM in improving the pricing performance of both its components. We can therefore price European and forward starting options to assess the overall effect of the proposed SLV mixing.

**European options**

We test the WSLVM performance with respect to short and long dated European call options. Figure 2.2 shows the results for maturities in the range \(\{1M, 3M, 1Y, 3Y\}\) in order to assess the performance over the entire term structure of volatility surface. As comparison we also priced the options within the pure LV and SV models. For the latter we consider both parameters sets obtained in Section 1.3.2 (i.e. with and without the constraint \(\beta \geq 1\)). In Appendix 2.A we report the numerical results in terms of difference between market and models implied volatilities. The following Table summarizes the evidences found in terms of mean absolute error with respect to market implied volatilities:

<table>
<thead>
<tr>
<th>Maturity</th>
<th>LV (\beta \geq 0)</th>
<th>WSSVM (\beta \geq 0)</th>
<th>WSSVM (\beta \geq 1)</th>
<th>WSLVM (\beta \geq 1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1M</td>
<td>0.0765</td>
<td>0.4138</td>
<td>1.8937</td>
<td>0.2115</td>
</tr>
<tr>
<td>3M</td>
<td>0.1945</td>
<td>0.7323</td>
<td>2.3794</td>
<td>0.1750</td>
</tr>
<tr>
<td>1Y</td>
<td>0.1329</td>
<td>0.1269</td>
<td>0.4075</td>
<td>0.1307</td>
</tr>
<tr>
<td>3Y</td>
<td>0.1720</td>
<td>0.1172</td>
<td>0.2099</td>
<td>0.1412</td>
</tr>
<tr>
<td>Overall</td>
<td>0.1474</td>
<td>0.3317</td>
<td>1.1405</td>
<td>0.1609</td>
</tr>
</tbody>
</table>

Table 2.1: Mean absolute error in volatility points \((1 = 1\% \text{ difference with respect to market values})\).

\(^2\)All the numerical tests performed in this Chapter have been implemented on the same machine used in Chapter 1. Also in this case, algorithms are coded in Matlab and then compiled as MEX files.
The conclusion is that if we enhance the WSVM with a LV term, then the constraint induced by the parameter $\beta$ can be fully neutralized: the marginal asset probability distribution in WSLVM substantially matches the market one. We see that the performance of WSLVM is comparable to the pure LV model. This is particularly evident if we focus on the shorter maturities: for options with strike prices far from the at-the-money level, the improvement over the SV parametrizations is quite substantial. The result is a direct consequence of the SLV mechanism: the compensation effect due to the leverage function fill the gap between market and SV model. For longer maturities, since the SV model already give adequate results, the adjustment due to the local
volatility term is less pronounced. However, also in this case that WSLVM gives a better fit of market values than the WSVM with $\beta \geq 1$. Computationally the affine nature of WSVM makes the pricing of European options quite fast without having to rely on simulations. However, the possibility to exploit new efficient simulation scheme for Wishart-based models allows to get prices in acceptable time also within the WSLVM. CPU time is then comparable to the ones shown in Table 1.1 for the pure SV case: the additional calculation time due to the presence of the local volatility term is indeed negligible.

Forward starting options

Having confirmed the ability of WSLVM in overcoming the limitations of pure WSVM with respect to the pricing of European claims we can now test if the new framework is also able to produce realistic forward volatility dynamics.

We price two different sets of forward starting options that starts, respectively, at $T_1 = 3$ months and $T_1 = 2$ years with maturity in one year. We can exploit the analytical tractability of WSVM to price forward starting options via FFT in the pure SV case. As shown in [34] the forward characteristic function of log-returns can be expressed in terms of the characteristic function of the Wishart process. This gives us an insight about the nature of this kind of options: they are basically pure variance derivatives. On the other hand, we need to rely on Monte Carlo simulations for the pricing with LV and WSLVM. Figure 2.3 illustrates the results in terms of Black-Scholes implied volatility, that is the value of $\sigma_{BS}$ such that

$$C_{FS_{Model}}(K, t_0, T_1, T_2) = e^{-rT_1}C_{BS}(1, K, T_2 - T_1, \sigma_{BS}),$$

where $C_{FS_{Model}}(K, t_0, T_1, T_2)$ is the model price of a forward starting option that starts at $T_1$ and expires at $T_2$ with strike $K$ while $C_{BS}(S_0, K, T, \sigma)$ is the Black-Scholes price of a plain vanilla call option. The message we get is clear: within the WSLVM framework we are still able to reproduce the proper volatility dynamics of the pure SV model. As we can see in Figure 2.3 as time goes by the forward implied volatility in the LV case becomes flatter showing the inadequacy of the model in describing how the volatility actually evolves. This does not happen for the SV and WSLVM where the skew persists through time. The ability of preserving a forward implied volatility shape similar to the SV case is a well-known feature of SLV models, as shown for example in [43] and [90] for the Heston Stochastic Local Volatility model. However, according to our knowledge, this is the first time that such a result is obtained for a multi-factor specification of the SV component.

2.4 The Wishart Stochastic Local Covariance model

In the previous Chapter we have shown that the reduced WASC model is able to generalize the Heston model in a multi-asset framework embedding a rich dependence
2.4. The Wishart Stochastic Local Covariance model

Figure 2.3: Implied volatility for forward starting call options on DAX index. \( T_1 = 3M \) (left), \( T_1 = 2Y \) (right). In both cases \( T_2 = T_1 + 1Y \).

structure among all the state variables involved. Unfortunately, in order to grant a realistic dynamics of Wishart generated cross-asset correlations, we found necessary to enforce the (very) restrictive condition of positive definiteness for the Wishart process. This in turn has the unpleasant effect to produce an extremely rigid implied volatility structure.

2.4.1 Model dynamics

With the objective of alleviating such a distortion, we now introduce the Wishart Stochastic Local Covariance Model (WSLCM), where the risk-neutral joint behavior of a \( d \)-dimensional vector of forward asset prices is

\[
df(t) = \text{diag} \begin{bmatrix} f(t) \end{bmatrix} \Psi[f(t), \Sigma(t)] \, db(t) \tag{2.18}
\]

with \( \Sigma(t) \) and \( b(t) \) defined, respectively, by (1.1) and (1.77). Here, the matrix function \( \Psi[f(t), \Sigma(t)] \) combines the local and the stochastic volatility components such that the instantaneous covariance matrix among log-assets is

\[
(dy(t)) (dy(t))^\top = \Psi \Psi^\top dt = \Psi^2 dt, \tag{2.19}
\]

with \( y(t) = [\log(f_1(t), f_2(t), ..., f_d(t))] \). A direct application of Theorem 2 guarantees that the single-asset terminal probability distributions induced by (2.18) are consistent with market ones provided that we have

\[
\mathbb{E}_Q^\otimes \left[ \Psi_{i,i}^2[f(t), \Sigma(t)] | f_i(t) = f \right] = \sigma_{LV,i}(t, f), \tag{2.20}
\]
for $i = 1, \ldots, d$ and $\sigma_{LV,i}$ is the local volatility calibrated to the $i$-th asset option prices. In order to fulfill this requirement, we assume the following parametrization for $\Psi$:

$$
\Psi[f(t), \Sigma(t)] = \text{diag} [\sigma(t, f(t))] \sqrt{\Sigma(t)}
$$

where $\sigma(t, f(t))$ is a vector-valued leverage function whose values are given by

$$
\sigma_i(t, f) = \frac{\sigma_{LV,i}(t, f)}{\sqrt{\mathbb{E}^Q[\Sigma_i(t)|f_i(t) = f]}}.
$$

It is worthwhile to point out that the choice of $\Psi$ is not unique and different specifications can accommodate alternative modelling features. If we consider, for example, the FX market, a desirable property of a pricing model would be to preserve the symmetry under inversion and triangulation of FX rates. In that case, matrix $\Psi$ in (2.21) is not the optimal choice since it would lead to mis-specified cross-rates dynamics. We refer to [38] for the treatment of the topic in a multi-asset hybrid extension of (a simplified) Heston model. Nonetheless, since our main concern here is to focus on equity markets, parametrization (2.21) represents an adequate setting. In particular, considering for simplicity $d = 2$, we have that the WSLCM asset instantaneous covariance matrix reads as

$$
(dy(t)) (dy(t))^\top = \Psi^2 dt = \begin{bmatrix}
\sigma_1^2 \Sigma_1(t) & \sigma_1 \sigma_2 \Sigma_{12}(t) \\
\sigma_1 \sigma_2 \Sigma_{12}(t) & \sigma_2^2 \Sigma_2(t)
\end{bmatrix},
$$

where we use the shorthand notation $\sigma_i = \sigma_i(t, e^{y_i(t)})$. The asset covariance is then hybrid (thus the name of the model), since it now depends on $y_1(t)$, $y_2(t)$ and $\Sigma_{12}(t)$. Furthermore, quite interestingly, we end up with the same correlation structure induced by the pure stochastic volatility model. From (2.23), indeed, it results that

$$
\text{Corr} [dy_1(t), dy_2(t)] = \frac{\Sigma_{12}(t)}{\sqrt{\Sigma_1(t) \Sigma_2(t)}} dt.
$$

This is due to the peculiar form of matrix $\Psi$. An inherent advantage of this result is that we are granted with cross-asset correlations that lie in the interval $[-1, 1]$ as soon as the Wishart process is well defined. For the sake of generality, though, we remark that the overall effects can be more subtle and would require a thorough analysis of the induced dependence structure that we leave to further research. In particular, it would be interesting to study the differences in terms of pricing of multi-asset securities with respect to hybrid specifications of asset correlation.

The framework proposed naturally applies to the most general parametrization of Wishart process. However, we consider a reduced form by setting matrix $M$ diagonal. In this setting, the WSLCM is able to match the prices of univariate path-dependent options generated with an hybrid Heston model whose parameters are set as illustrated in Section 1.4.1. Indeed, as already pointed out for the WASC model, it is possible to show
2.4. The Wishart Stochastic Local Covariance model

that each asset dynamics is equivalent to an individual instance of the hybrid Heston model. This allows, for example, to rely on existing techniques specifically devised for the Heston case, such as the numerical schemes for the solution of the Fokker-Planck PDE in the calibration of $\sigma_i(t, f)$. More importantly, in the light of the above, we can interpret the WSLCM as an intuitive multi-asset extension of the Heston model able to introduce an high degree of flexibility in designing the dependence structure thanks to the underlying Wishart process.

2.4.2 Numerical Results

This section is devoted to illustrate a realistic implementation of the WSLCM. We consider here a two-assets specification of the model intended to price derivatives on the pair of indices EuroStoxx50-DAX. We refer to section 2.2.1 for the calibration of the two non parametric local volatility surfaces (one for each asset). Additionally, we consider the Wishart-based parameters in the third column of Table 1.7 i.e. we assume to deal with the parameters set obtained by requiring the Wishart process to be defined in the interior of the cone $S_n^+$ ($\mathbb{R}$). This setting is motivated by the fact that our goal is to construct a model able to produce realistic cross-asset correlations.

Having specified the LV and SV components, we can now address the task of calibrating the bivariate leverage function. To this extent, we estimate (2.22) by means of simulation based approach (2.11) coupled with the GVA scheme devised in Section 1.4.4. This scheme can be extended to deal with the discretization of WSLCM trajectories by simply modifying (1.102) into

$$\hat{y}_i(t + \Delta) = \hat{y}_i(t) - \frac{1}{2} \sigma_i^2(t, e^{\hat{y}_i(t)}) \hat{S}_i(t) \Delta + \sigma_i(t, e^{\hat{y}_i(t)}) \sqrt{\hat{S}_i(t)} \sum_{j=1}^{d+i} \hat{\ell}_{d+i,j}(t) \hat{w}_j(t) \sqrt{\hat{S}_i(t)} \hat{w}_j(t)$$

for $i = 1, ..., d$, where all other terms remain unchanged.

Despite the evidences found in Chapter 1, indeed, it is our opinion that the GVA scheme should be preferred over the simpler TE scheme for calibration purposes. This is due to the fact that, as already stressed, GVA scheme allows for an accurate discretization of Wishart process. On the contrary, within the TE scheme we cannot bound the process $\Sigma(t)$ to remain in the cone of positive semi definite matrices and diagonal elements can also become negative with non zero probability. Even worse, the truncation mechanism could severely affect the estimate of (conditional) expected values in (2.22) and, ultimately, lead to wrong calibrated leverage functions. Having this in mind, we perform the calibration of leverage function by sampling $5 \times 10^5$ trajectories on a daily time grid. As in the single-asset case, we consider 50 equally weighted bins at each time step. We use the model so defined to price European call options written on each of the two indices, with maturity of one month and three years. As comparison, we price the basket of options in a pure LV setting and in the WASC model for which we consider
Chapter 2. A New Class of Multidimensional Wishart-based hybrid models

the unconstrained parameters and those obtained when the condition \( \beta \geq 3 \) is satisfied (respectively given in the first and third column of Table 1.7). In Appendix 2.B we exhibit the full set of numerical results. We also report the mean absolute error in volatility terms with respect to the market:

<table>
<thead>
<tr>
<th>Asset</th>
<th>Maturity</th>
<th>LV</th>
<th>WASC ( \beta \geq 0 )</th>
<th>WASC ( \beta \geq 3 )</th>
<th>WSLCM</th>
</tr>
</thead>
<tbody>
<tr>
<td>EuroStoxx50</td>
<td>1M</td>
<td>0.2078</td>
<td>1.2459</td>
<td>3.2357</td>
<td>0.3357</td>
</tr>
<tr>
<td></td>
<td>3Y</td>
<td>0.0894</td>
<td>0.5197</td>
<td>0.7123</td>
<td>0.1301</td>
</tr>
<tr>
<td>DAX</td>
<td>1M</td>
<td>0.0765</td>
<td>0.8727</td>
<td>2.9537</td>
<td>0.3818</td>
</tr>
<tr>
<td></td>
<td>3Y</td>
<td>0.1720</td>
<td>0.3108</td>
<td>0.4866</td>
<td>0.0695</td>
</tr>
<tr>
<td>Overall</td>
<td></td>
<td>0.1356</td>
<td>0.6885</td>
<td>1.6580</td>
<td>0.2097</td>
</tr>
</tbody>
</table>

Table 2.2: Mean absolute error in volatility points (1% difference with respect to market values).

The most evident result is that the WSLCVM is able to overcome the limitations due to the request of a positive definite Wishart process. Not only the improvement with respect to the constrained WASC parametrization is dramatic (the average volatility error drops from 1.6580% to 0.2097%), but also the comparison with the fully calibrated SV model highlights a consistent overperformance of the new hybrid setting (the error is reduced by about 70% on average). This effect can be appreciated by looking at Figure 2.4 where the models induced implied volatilities are shown together with market ones. In particular, we want to stress that the only difference between the WSLCM and and the WASC with the condition \( \beta \geq 3 \) is the presence of the additional local adjustment factor in the former dynamics. As a consequence of (2.22), the hybrid model is in line with the LV one and reproduces satisfactorily the market implied volatility smiles, even when short times to maturity are considered. Overall, the new model succeeds in fitting the market evidences on single asset markets and allows a sound modelling of the dependence structure (we refer to the left panel of Figure 1.11 for an illustration of induced trajectories of cross-asset correlation).

2.5 Concluding remarks

In this Chapter, we built a new class of hybrid models with the remarkable feature of embedding a matrix-defined dynamics for the stochastic evolution of variance factors. Numerical tests highlight that the new models effectively incorporate the advantages of LV approach in reproducing plain vanilla market evidences, still preserving the flexibility of Wishart-based pure SV models. In our framework, the contribution of the additional LV component is even more relevant if we take into account the fact that
2.5. Concluding remarks

Figure 2.4: Implied volatility for European call options for maturities $T = 1$M (left) and $T = 3$Y (right). Upper panels: EuroStoxx50 index. Lower panels: DAX index.

Market calibrated parameters usually violate the condition for existence and uniqueness of (weak) solution to the Wishart SDE. We showed that the proposed SLV mixing presents an effective, probably unique, tool to deal with this problem within the class of Wishart-based pricing models. Additionally, WSLVM is found to be adequate even when we study the resulting volatility dynamics: it preserves the shape of forward implied volatility originated by the pure multidimensional SV model. This feature, in conjunction with the ability of properly manage the time structure of the skew, makes the WSLVM a valid alternative to price forward implied volatility dependent payoffs. Further, in the light of the evidences shown in Chapter 1, it seems appropriate to consider WSLCM as the most genuine and comprehensive framework proposed so far to
Chapter 2. A New Class of Multidimensional Wishart-based hybrid models

extend the famous Heston model to the multi-asset case. The resulting model, indeed, is able to generate a sophisticated dependence structure among assets and variance factors. More importantly, the new model succeeds in eliminating any kind of mispricing of plain vanilla claims due to the stronger condition that we need to impose on model parameters in order to get reasonable cross-asset correlations dynamics.

Calibration and pricing in the new framework are easily dealt with thanks to the numerical techniques proposed in Chapter 1. In particular, we extensively rely on the simulation schemes devised for WSVM and WASC that allow to exploit the exact sampling of Wishart process in [3].

In our opinion the new models could turn out to be a comprehensive modelling framework to price a heterogeneous range of exotic derivatives consistently with European claims when a flexible multidimensional dynamics of variance factors is required.
### 2.A WSLVM: Numerical Results

<table>
<thead>
<tr>
<th>Time</th>
<th>$K/f_0$ (%)</th>
<th>Market</th>
<th>Error</th>
<th>LV</th>
<th>WSVM $\beta \geq 0$</th>
<th>WSVM $\beta \geq 1$</th>
<th>WSLVM</th>
</tr>
</thead>
<tbody>
<tr>
<td>75</td>
<td>40.96</td>
<td>-0.12</td>
<td>0.14</td>
<td>5.35</td>
<td>1.32</td>
<td></td>
<td></td>
</tr>
<tr>
<td>80</td>
<td>38.25</td>
<td>-0.03</td>
<td>0.24</td>
<td>4.26</td>
<td>0.62</td>
<td></td>
<td></td>
</tr>
<tr>
<td>85</td>
<td>35.56</td>
<td>0.00</td>
<td>0.36</td>
<td>3.20</td>
<td>0.26</td>
<td>0.00</td>
<td>0.26</td>
</tr>
<tr>
<td>90</td>
<td>32.89</td>
<td>-0.03</td>
<td>0.46</td>
<td>2.17</td>
<td>0.08</td>
<td>0.08</td>
<td></td>
</tr>
<tr>
<td>95</td>
<td>30.25</td>
<td>-0.03</td>
<td>0.49</td>
<td>1.19</td>
<td>0.02</td>
<td>0.02</td>
<td></td>
</tr>
<tr>
<td>97.5</td>
<td>28.95</td>
<td>-0.03</td>
<td>0.46</td>
<td>0.72</td>
<td>0.00</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td>1 M</td>
<td>75</td>
<td>38.66</td>
<td>0.17</td>
<td>1.21</td>
<td>-0.46</td>
<td>-0.12</td>
<td>-0.12</td>
</tr>
<tr>
<td>100</td>
<td>70</td>
<td>36.09</td>
<td>0.22</td>
<td>-0.11</td>
<td>2.67</td>
<td>0.07</td>
<td></td>
</tr>
<tr>
<td>102.5</td>
<td>70</td>
<td>33.54</td>
<td>0.10</td>
<td>-0.08</td>
<td>1.73</td>
<td>-0.03</td>
<td></td>
</tr>
<tr>
<td>105</td>
<td>70</td>
<td>31.02</td>
<td>0.04</td>
<td>-0.12</td>
<td>0.83</td>
<td>-0.09</td>
<td></td>
</tr>
<tr>
<td>110</td>
<td>70</td>
<td>28.55</td>
<td>0.01</td>
<td>-0.22</td>
<td>-0.01</td>
<td>-0.10</td>
<td></td>
</tr>
<tr>
<td>115</td>
<td>70</td>
<td>27.34</td>
<td>0.00</td>
<td>-0.28</td>
<td>-0.39</td>
<td>-0.10</td>
<td></td>
</tr>
<tr>
<td>120</td>
<td>70</td>
<td>26.18</td>
<td>-0.02</td>
<td>-0.35</td>
<td>-0.73</td>
<td>-0.11</td>
<td></td>
</tr>
<tr>
<td>125</td>
<td>70</td>
<td>25.07</td>
<td>-0.05</td>
<td>-0.41</td>
<td>-1.02</td>
<td>-0.15</td>
<td></td>
</tr>
<tr>
<td>130</td>
<td>70</td>
<td>24.04</td>
<td>-0.09</td>
<td>-0.45</td>
<td>-1.25</td>
<td>-0.20</td>
<td></td>
</tr>
<tr>
<td>105</td>
<td>70</td>
<td>22.37</td>
<td>-0.09</td>
<td>-0.47</td>
<td>-1.42</td>
<td>-0.19</td>
<td></td>
</tr>
<tr>
<td>110</td>
<td>70</td>
<td>21.49</td>
<td>-0.05</td>
<td>-0.41</td>
<td>-1.05</td>
<td>-0.14</td>
<td></td>
</tr>
<tr>
<td>115</td>
<td>70</td>
<td>21.60</td>
<td>-0.02</td>
<td>-0.20</td>
<td>-0.04</td>
<td>-0.13</td>
<td></td>
</tr>
<tr>
<td>120</td>
<td>70</td>
<td>22.56</td>
<td>0.07</td>
<td>0.33</td>
<td>1.47</td>
<td>-0.08</td>
<td></td>
</tr>
<tr>
<td>125</td>
<td>70</td>
<td>24.01</td>
<td>0.19</td>
<td>1.12</td>
<td>3.22</td>
<td>0.01</td>
<td></td>
</tr>
<tr>
<td>130</td>
<td>70</td>
<td>27.35</td>
<td>0.38</td>
<td>2.91</td>
<td>6.67</td>
<td>0.20</td>
<td></td>
</tr>
<tr>
<td>140</td>
<td>70</td>
<td>30.57</td>
<td>0.58</td>
<td>4.59</td>
<td>9.69</td>
<td>0.40</td>
<td></td>
</tr>
</tbody>
</table>

Table 2.3: Results of numerical experiments: Part I. Market indicates the implied volatility observed on 3rd February 2016 for a given call option on DAX index. The Error columns contain the difference between market implied volatility and the one obtained with the corresponding model.
| Time | \( K/f_0 \) | Market | Error | \( \text{LV} \ | \beta \geq 0 \) | \( \text{WSVM} \ | \beta \geq 1 \) | WSLVM |
|------|-------------|--------|------|-------------------|------------------|------|
| 50   | 40.53       | -0.62  | 0.01 | 2.16              | 1.15             |
| 60   | 36.71       | -0.35  | 0.07 | 1.47              | 0.45             |
| 70   | 33.20       | -0.19  | 0.05 | 0.91              | 0.18             |
| 75   | 31.55       | -0.16  | 0.03 | 0.68              | 0.10             |
| 80   | 29.95       | -0.13  | 0.01 | 0.47              | 0.05             |
| 85   | 28.42       | -0.11  | 0.00 | 0.30              | 0.01             |
| 90   | 26.94       | -0.10  | 0.02 | 0.15              | -0.01            |
| 95   | 25.53       | -0.10  | 0.05 | 0.04              | -0.03            |
| 97.5 | 24.86       | -0.09  | 0.08 | -0.01             | -0.03            |
| 100  | 24.21       | -0.09  | 0.12 | -0.05             | -0.03            |
| 102.5| 23.58       | -0.09  | 0.16 | -0.09             | -0.03            |
| 105  | 22.99       | -0.10  | 0.21 | -0.11             | -0.04            |
| 110  | 21.90       | -0.09  | 0.30 | -0.15             | -0.04            |
| 115  | 20.96       | -0.07  | 0.36 | -0.16             | -0.02            |
| 120  | 20.19       | -0.06  | 0.34 | -0.16             | -0.03            |
| 125  | 19.61       | -0.06  | 0.25 | -0.12             | -0.04            |
| 130  | 19.22       | -0.04  | 0.11 | -0.06             | -0.04            |
| 140  | 18.92       | 0.00   | -0.11| 0.15              | -0.09            |
| 150  | 19.10       | 0.06   | -0.11| 0.48              | -0.12            |
| 50   | 31.24       | -0.52  | -0.18| 0.57              | 0.65             |
| 60   | 29.11       | -0.33  | -0.09| 0.47              | 0.38             |
| 70   | 27.22       | -0.23  | -0.05| 0.34              | 0.22             |
| 75   | 26.34       | -0.21  | -0.04| 0.27              | 0.17             |
| 80   | 25.50       | -0.19  | -0.05| 0.19              | 0.14             |
| 85   | 24.71       | -0.17  | -0.06| 0.11              | 0.12             |
| 90   | 23.96       | -0.16  | -0.08| 0.04              | 0.12             |
| 95   | 23.24       | -0.15  | -0.10| -0.04             | 0.12             |
| 97.5 | 22.90       | -0.15  | -0.12| -0.08             | 0.11             |
| 100  | 22.57       | -0.14  | -0.13| -0.11             | 0.11             |
| 102.5| 22.26       | -0.14  | -0.14| -0.15             | 0.10             |
| 105  | 21.95       | -0.14  | -0.15| -0.18             | 0.10             |
| 110  | 21.38       | -0.13  | -0.17| -0.22             | 0.08             |
| 115  | 20.86       | -0.12  | -0.19| -0.25             | 0.07             |
| 120  | 20.40       | -0.11  | -0.19| -0.26             | 0.06             |
| 125  | 20.00       | -0.11  | -0.18| -0.24             | 0.04             |
| 130  | 19.67       | -0.10  | -0.15| -0.19             | 0.03             |
| 140  | 19.18       | -0.09  | -0.04| 0.00              | 0.02             |
| 150  | 18.91       | -0.07  | 0.11 | 0.28              | 0.04             |

Table 2.4: Results of numerical experiments: Part II. Market indicates the implied volatility observed on 3rd February 2016 for a given call option on DAX index. The Error columns contain the difference between market implied volatility and the one obtained with the corresponding model.
### 2.B WSLCM: Numerical Results

<table>
<thead>
<tr>
<th>Time</th>
<th>$K/f_0$ (%)</th>
<th>Market</th>
<th>Error</th>
<th>LV</th>
<th>WASC $\beta \geq 0$</th>
<th>WASC $\beta \geq 1$</th>
<th>WSLCM</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>75</td>
<td>40.18</td>
<td>0.49</td>
<td>3.77</td>
<td>11.86</td>
<td>-1.56</td>
<td></td>
</tr>
<tr>
<td>1 M</td>
<td>80</td>
<td>36.90</td>
<td>0.31</td>
<td>2.78</td>
<td>9.63</td>
<td>-0.62</td>
<td></td>
</tr>
<tr>
<td></td>
<td>85</td>
<td>33.58</td>
<td>0.18</td>
<td>1.81</td>
<td>7.34</td>
<td>-0.19</td>
<td></td>
</tr>
<tr>
<td></td>
<td>90</td>
<td>30.21</td>
<td>0.08</td>
<td>0.89</td>
<td>5.00</td>
<td>-0.05</td>
<td></td>
</tr>
<tr>
<td></td>
<td>95</td>
<td>26.83</td>
<td>0.03</td>
<td>0.08</td>
<td>2.64</td>
<td>0.05</td>
<td></td>
</tr>
<tr>
<td></td>
<td>97.5</td>
<td>25.19</td>
<td>0.02</td>
<td>-0.24</td>
<td>1.50</td>
<td>0.08</td>
<td></td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>23.63</td>
<td>0.00</td>
<td>-0.48</td>
<td>0.44</td>
<td>0.09</td>
<td></td>
</tr>
<tr>
<td></td>
<td>102.5</td>
<td>22.22</td>
<td>-0.05</td>
<td>-0.63</td>
<td>-0.47</td>
<td>0.05</td>
<td></td>
</tr>
<tr>
<td></td>
<td>105</td>
<td>21.05</td>
<td>-0.10</td>
<td>-0.70</td>
<td>-1.15</td>
<td>-0.02</td>
<td></td>
</tr>
<tr>
<td></td>
<td>110</td>
<td>19.55</td>
<td>-0.03</td>
<td>-0.82</td>
<td>-1.69</td>
<td>0.04</td>
<td></td>
</tr>
<tr>
<td></td>
<td>115</td>
<td>19.03</td>
<td>0.10</td>
<td>-1.03</td>
<td>-1.31</td>
<td>0.12</td>
<td></td>
</tr>
<tr>
<td></td>
<td>120</td>
<td>19.06</td>
<td>0.28</td>
<td>-1.24</td>
<td>-0.49</td>
<td>0.41</td>
<td></td>
</tr>
<tr>
<td></td>
<td>125</td>
<td>19.34</td>
<td>0.49</td>
<td>-1.42</td>
<td>0.45</td>
<td>0.63</td>
<td></td>
</tr>
<tr>
<td></td>
<td>130</td>
<td>19.74</td>
<td>0.74</td>
<td>-1.56</td>
<td>1.34</td>
<td>0.80</td>
<td></td>
</tr>
<tr>
<td>3 Y</td>
<td>50</td>
<td>27.42</td>
<td>-0.22</td>
<td>0.79</td>
<td>1.99</td>
<td>-0.96</td>
<td></td>
</tr>
<tr>
<td></td>
<td>60</td>
<td>25.57</td>
<td>-0.16</td>
<td>0.61</td>
<td>1.55</td>
<td>-0.35</td>
<td></td>
</tr>
<tr>
<td></td>
<td>70</td>
<td>23.91</td>
<td>-0.13</td>
<td>0.39</td>
<td>1.12</td>
<td>-0.07</td>
<td></td>
</tr>
<tr>
<td></td>
<td>75</td>
<td>23.13</td>
<td>-0.12</td>
<td>0.27</td>
<td>0.90</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td></td>
<td>80</td>
<td>22.38</td>
<td>-0.11</td>
<td>0.14</td>
<td>0.69</td>
<td>0.05</td>
<td></td>
</tr>
<tr>
<td></td>
<td>85</td>
<td>21.65</td>
<td>-0.10</td>
<td>0.00</td>
<td>0.48</td>
<td>0.08</td>
<td></td>
</tr>
<tr>
<td></td>
<td>90</td>
<td>20.96</td>
<td>-0.09</td>
<td>-0.14</td>
<td>0.27</td>
<td>0.10</td>
<td></td>
</tr>
<tr>
<td></td>
<td>95</td>
<td>20.29</td>
<td>-0.08</td>
<td>-0.29</td>
<td>0.07</td>
<td>0.10</td>
<td></td>
</tr>
<tr>
<td></td>
<td>97.5</td>
<td>19.98</td>
<td>-0.08</td>
<td>-0.37</td>
<td>-0.03</td>
<td>0.09</td>
<td></td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>19.67</td>
<td>-0.08</td>
<td>-0.44</td>
<td>-0.12</td>
<td>0.09</td>
<td></td>
</tr>
<tr>
<td></td>
<td>102.5</td>
<td>19.37</td>
<td>-0.08</td>
<td>-0.51</td>
<td>-0.21</td>
<td>0.08</td>
<td></td>
</tr>
<tr>
<td></td>
<td>105</td>
<td>19.09</td>
<td>-0.08</td>
<td>-0.57</td>
<td>-0.28</td>
<td>0.07</td>
<td></td>
</tr>
<tr>
<td></td>
<td>110</td>
<td>18.60</td>
<td>-0.07</td>
<td>-0.64</td>
<td>-0.38</td>
<td>0.04</td>
<td></td>
</tr>
<tr>
<td></td>
<td>115</td>
<td>18.24</td>
<td>-0.06</td>
<td>-0.61</td>
<td>-0.36</td>
<td>0.01</td>
<td></td>
</tr>
<tr>
<td></td>
<td>120</td>
<td>18.07</td>
<td>-0.05</td>
<td>-0.42</td>
<td>-0.17</td>
<td>-0.02</td>
<td></td>
</tr>
<tr>
<td></td>
<td>125</td>
<td>18.08</td>
<td>-0.05</td>
<td>-0.09</td>
<td>0.18</td>
<td>-0.05</td>
<td></td>
</tr>
<tr>
<td></td>
<td>130</td>
<td>18.20</td>
<td>-0.04</td>
<td>0.33</td>
<td>0.62</td>
<td>-0.07</td>
<td></td>
</tr>
<tr>
<td></td>
<td>140</td>
<td>18.59</td>
<td>-0.04</td>
<td>1.22</td>
<td>1.59</td>
<td>-0.11</td>
<td></td>
</tr>
<tr>
<td></td>
<td>150</td>
<td>19.02</td>
<td>-0.03</td>
<td>2.05</td>
<td>2.53</td>
<td>-0.14</td>
<td></td>
</tr>
</tbody>
</table>

**Table 2.5:** Results of numerical experiments: Part I. Market indicates the implied volatility observed on 3rd February 2016 for a given call option on EuroStoxx50 index. The Error columns contain the difference between market implied volatility and the one obtained with the corresponding model.
### Table 2.6: Results of numerical experiments: Part II

Market indicates the implied volatility observed on 3rd February 2016 for a given call option on DAX index. The Error columns contain the difference between market implied volatility and the one obtained with the corresponding model.

<table>
<thead>
<tr>
<th>Time</th>
<th>$K/f_0$ (%)</th>
<th>Market</th>
<th>Error LV</th>
<th>WASC $\beta \geq 0$</th>
<th>WASC $\beta \geq 1$</th>
<th>WSCLCM</th>
</tr>
</thead>
<tbody>
<tr>
<td>75</td>
<td>40.96</td>
<td>-0.12</td>
<td>1.76</td>
<td>6.60</td>
<td>-0.89</td>
<td></td>
</tr>
<tr>
<td>80</td>
<td>38.25</td>
<td>-0.03</td>
<td>1.17</td>
<td>5.23</td>
<td>-0.29</td>
<td></td>
</tr>
<tr>
<td>85</td>
<td>35.56</td>
<td>0.00</td>
<td>0.62</td>
<td>3.90</td>
<td>-0.05</td>
<td></td>
</tr>
<tr>
<td>90</td>
<td>32.89</td>
<td>-0.03</td>
<td>0.14</td>
<td>2.60</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td>95</td>
<td>30.25</td>
<td>-0.03</td>
<td>-0.24</td>
<td>1.36</td>
<td>0.03</td>
<td></td>
</tr>
<tr>
<td>97.5</td>
<td>28.95</td>
<td>-0.03</td>
<td>-0.39</td>
<td>0.78</td>
<td>0.04</td>
<td></td>
</tr>
<tr>
<td>1 M</td>
<td>100</td>
<td>27.68</td>
<td>-0.05</td>
<td>-0.51</td>
<td>0.24</td>
<td>0.03</td>
</tr>
<tr>
<td>102.5</td>
<td>26.45</td>
<td>-0.10</td>
<td>-0.59</td>
<td>-0.23</td>
<td>-0.02</td>
<td></td>
</tr>
<tr>
<td>105</td>
<td>25.28</td>
<td>-0.17</td>
<td>-0.65</td>
<td>-0.62</td>
<td>-0.10</td>
<td></td>
</tr>
<tr>
<td>110</td>
<td>23.3</td>
<td>-0.16</td>
<td>-0.80</td>
<td>-0.96</td>
<td>-0.11</td>
<td></td>
</tr>
<tr>
<td>115</td>
<td>22.14</td>
<td>-0.08</td>
<td>-0.86</td>
<td>-0.29</td>
<td>-0.04</td>
<td></td>
</tr>
<tr>
<td>120</td>
<td>22.26</td>
<td>-0.08</td>
<td>-0.32</td>
<td>1.92</td>
<td>0.22</td>
<td></td>
</tr>
<tr>
<td>125</td>
<td>23.64</td>
<td>-0.01</td>
<td>1.10</td>
<td>5.84</td>
<td>1.23</td>
<td></td>
</tr>
<tr>
<td>130</td>
<td>25.78</td>
<td>0.18</td>
<td>3.07</td>
<td>10.80</td>
<td>2.29</td>
<td></td>
</tr>
<tr>
<td>50</td>
<td>31.24</td>
<td>-0.52</td>
<td>0.28</td>
<td>1.74</td>
<td>-0.24</td>
<td></td>
</tr>
<tr>
<td>60</td>
<td>29.11</td>
<td>-0.33</td>
<td>0.17</td>
<td>1.20</td>
<td>-0.06</td>
<td></td>
</tr>
<tr>
<td>70</td>
<td>27.22</td>
<td>-0.23</td>
<td>0.03</td>
<td>0.72</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td>75</td>
<td>26.34</td>
<td>-0.21</td>
<td>-0.04</td>
<td>0.50</td>
<td>0.01</td>
<td></td>
</tr>
<tr>
<td>80</td>
<td>25.5</td>
<td>-0.19</td>
<td>-0.12</td>
<td>0.30</td>
<td>0.01</td>
<td></td>
</tr>
<tr>
<td>85</td>
<td>24.71</td>
<td>-0.17</td>
<td>-0.19</td>
<td>0.11</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td>3 Y</td>
<td>90</td>
<td>23.96</td>
<td>-0.16</td>
<td>-0.27</td>
<td>-0.06</td>
<td>0.00</td>
</tr>
<tr>
<td>95</td>
<td>23.24</td>
<td>-0.15</td>
<td>-0.34</td>
<td>-0.20</td>
<td>-0.02</td>
<td></td>
</tr>
<tr>
<td>97.5</td>
<td>22.9</td>
<td>-0.15</td>
<td>-0.37</td>
<td>-0.27</td>
<td>-0.02</td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>22.57</td>
<td>-0.14</td>
<td>-0.40</td>
<td>-0.32</td>
<td>-0.03</td>
<td></td>
</tr>
<tr>
<td>102.5</td>
<td>22.26</td>
<td>-0.14</td>
<td>-0.43</td>
<td>-0.37</td>
<td>-0.04</td>
<td></td>
</tr>
<tr>
<td>105</td>
<td>21.95</td>
<td>-0.14</td>
<td>-0.46</td>
<td>-0.41</td>
<td>-0.05</td>
<td></td>
</tr>
<tr>
<td>110</td>
<td>21.38</td>
<td>-0.13</td>
<td>-0.50</td>
<td>-0.46</td>
<td>-0.06</td>
<td></td>
</tr>
<tr>
<td>115</td>
<td>20.86</td>
<td>-0.12</td>
<td>-0.53</td>
<td>-0.47</td>
<td>-0.08</td>
<td></td>
</tr>
<tr>
<td>120</td>
<td>20.40</td>
<td>-0.11</td>
<td>-0.53</td>
<td>-0.44</td>
<td>-0.10</td>
<td></td>
</tr>
<tr>
<td>125</td>
<td>20.00</td>
<td>-0.11</td>
<td>-0.50</td>
<td>-0.34</td>
<td>-0.12</td>
<td></td>
</tr>
<tr>
<td>130</td>
<td>19.67</td>
<td>-0.10</td>
<td>-0.45</td>
<td>-0.20</td>
<td>-0.14</td>
<td></td>
</tr>
<tr>
<td>140</td>
<td>19.18</td>
<td>-0.09</td>
<td>-0.27</td>
<td>0.25</td>
<td>-0.16</td>
<td></td>
</tr>
<tr>
<td>150</td>
<td>18.91</td>
<td>-0.07</td>
<td>-0.02</td>
<td>0.88</td>
<td>-0.18</td>
<td></td>
</tr>
</tbody>
</table>
Part II

On the Hull-White model for the computation of CVA with WWR
On a simple and effective methodology to deal with the computation of CVA with WWR within the Hull-White Model

3.1 Introduction

As a direct consequence of 2008 financial turmoil, counterparty credit risk has become substantial in OTC derivatives transactions. In particular the Credit Value Adjustment (CVA) is meant to measure the impact of counterparty riskiness on a derivative portfolio value as requested by the current Basel III regulatory framework. Accounting standards (IFRS 13, FAS 157), moreover, require a CVA\textsuperscript{2} adjustment as part of a consistent fair value measurement of financial instruments. In the following we focus on the so-called unilateral CVA, i.e. the party that carries out the valuation is considered default-free. Even if restrictive, this is the underlying assumption for counterparty risk valuation in the Capital Requirement Regulation (CRR, Art. 381). Unilateral CVA has also played a crucial role in the recent Asset Quality Review performed by the ECB\textsuperscript{44}. For

\textsuperscript{1}This Chapter is based on two articles jointly written with Roberto Baviera (Politecnico di Milano) and Paolo Pellicioli: \textit{A note on CVA and wrong way risk} (International Journal of Financial Engineering, 3(2), 2016) and \textit{CVA with Wrong Way Risk in the Presence of Early Exercise} (to appear in Innovations in Derivatives Markets - Fixed Income Modeling, Valuation Adjustments, Risk Management, and Regulation, Springer, 2016)

\textsuperscript{2}Even if in this paper we focus on CVA pricing, it is worthwhile to note that accounting standards ask also for a Debt Value Adjustment (DVA) to take into account the own credit risk.
Chapter 3. On the computation of CVA with WWR within the Hull-White Model

a detailed discussion on other forms of CVA we refer the interested reader to [57]. CVA is strongly affected by derivative transaction arrangements: exposure depends on collateral and netting agreement between the two counterparties that have written the derivative contracts of interest. Despite the increased use of collateral, however, a significant portion of OTC derivatives remains uncollateralized. This is mainly due to the nature of the counterparties involved, such as corporates and sovereigns, without the liquidity and operational capacity to adhere to daily collateral calls. In such cases, an institution must consider the impact of counterparty risk on the overall portfolio value and a correct CVA quantification acquires even more importance. A wide literature has been produced on the topic in recent years as for example [14] and [57] that give a comprehensive overview of CVA computation and the more general topic of Counterparty Credit Risk management. It seems, however, that attention has been mainly paid to CVA with respect to portfolios of European style derivatives. Dealing with derivatives with early exercise features is even more delicate. In these cases, indeed, we need to take also into account the exercise strategy and the fact that exposure falls to zero after the exercise.

A peculiar problem that we encounter in CVA computation is the presence of the so-called Wrong Way Risk (WWR), that is the non negligible dependency between the value of a derivatives portfolio and counterparty default probability. In particular we face WWR if a deterioration in counterparty creditworthiness is more likely when portfolio exposure increases. Several attempts have been made to deal with WWR. From a regulatory point of view, the Basel III Committee currently requires to correct by a multiplicative factor $\alpha = 1.4$ the CVA computed under hypothesis of market-credit independence. In this way the impact of WWR is considered equivalent to a 40% increase in standard CVA. However the Committee leaves room for financial institutions with approved models to apply for lower multipliers (floored at 1.2). This opportunity opens the way for more sophisticated models in order to reach a more efficient risk capital allocation. Relevant contributions on alternative approaches to manage WWR include copula-based modelling as in [21], introduction of jumps at default as in [77], the structural framework based on correlated Lévy processes in [7] and the stochastic hazard rate approach in [67]. In particular [67] introduces the idea to link the counterparty hazard rate to the portfolio value by means of an arbitrary monotone function (from now on we refer to this approach as the Hull-White model for CVA pricing in presence of WWR). The dependence structure in the Hull-White setting is described uniquely by one parameter that controls the impact of exposures on the hazard rate. Additionally a deterministic time dependent function is introduced to match the counterparty credit term structure observed on the market. In this framework CVA pricing in the presence of WWR involves just a small adjustment to the pricing machinery already in place in financial institutions. We only need to take into account the randomness incorporated into the counterparty default probabilities by means of the stochastic hazard rate and
3.2. CVA pricing and WWR

price CVA with standard techniques. This is probably the most relevant property of the model: as soon as we associate a WWR parameter to a given counterparty-portfolio combination we are able to deal with WWR using the same pricing engine underlying standard CVA computation. As pointed out in [85], leveraging as much as possible on existing platforms should be one of the principles an optimal risk model should be shaped on.

However, the original approach in [67] relies on a Monte Carlo-based technique to determine the auxiliary deterministic function in order to calibrate the model on the counterparty credit structure. Obtaining this auxiliary function is the trickiest part in the calibration procedure, because it involves a “delicate” path-dependent problem that is difficult to implement for realistic portfolios. Here we show how it is possible to overcome such a limitation by transforming the path-dependent problem into a recursive one with a considerable reduction in the overall computational complexity. The basic idea is to consider discrete market factor dynamics and induce a change of probability such that the new set of (transition) probabilities are computed recursively in time. We present a straightforward implementation of our approach via tree methods. Trees are also a straightforward and well understood tool to manage the early termination in derivatives pricing. In the light of above, we propose to combine the new recursive formula and the tree-based dynamic programming in order to devise a simple and effective procedure to price CVA with WWR when American or Bermudan features are considered. The Chapter is organized as follows: in Section 3.2 we review the notion of CVA and illustrate the Hull-White approach to WWR modelling. In Section 3.3 we introduce the recursive algorithm meant to facilitate the implementation of Hull-White model. In Section 3.4 we propose an application of the novel technique in the context of CVA pricing with WWR in case of linear derivatives. In Section 3.5 we extend our methodology to American and Bermudan options and analyse the effects of early termination on CVA figures. Finally, in Section 3.6 we report some final remarks.

3.2 CVA pricing and WWR

For a given derivatives portfolio we can define the unilateral CVA as the risk-neutral expectation of the discounted loss that can be suffered over a given period of time

\[ CVA = (1 - R) \int_{t_0}^{T} B(t_0, t) \EE(t) PD(dt) , \]

where usually \( t_0 \) is the value date (hereinafter we set \( t_0 = 0 \) if not stated otherwise) and \( T \) is the longest maturity date in the portfolio. Here \( R \) is the Recovery rate, \( PD(dt) \) is the probability density of counterparty default between \( t \) and \( t + dt \) (with no default before \( t \)), and \( B(t_0, t) \EE(t) \) is the discounted expected exposure in \( t \). If interest rates are stochastic, the expected exposure is defined

\[ B(t_0, t) \EE(t) \equiv \EE[D(t_0, t) E(t)] , \]
Chapter 3. On the computation of CVA with WWR within the Hull-White Model

with $\mathbb{E}[:]$ the expectation operator given the information at value date $t_0$, $D(t_0, t)$ the stochastic discount and $E(t)$ the (stochastic) exposure at time $t$. The latter is inherently defined by the collateral agreement that the parties have in place: for example in uncollateralized transactions, $E(t)$ is simply the max w.r.t. zero of $v(t)$, the portfolio value at time $t$. For practical computation, the integral in (3.1) is approximated by choosing a discretized set of times $\mathcal{T} = \{t_i\}_{i=0,\ldots,n}$ with $t_n = T$. Within Basel III regulatory framework for financial institutions with an Internal Model Method, the standard approach for CVA valuation is

$$CVA = (1 - R) \sum_{i=1}^{n} B_i EE_i + B_{i-1} EE_{i-1}^2 PD_i ,$$

(3.2)

with $B_i$ that stands for $B(t_0, t_i)$ and

$$PD_i \equiv SP_{i-1} - SP_i ,$$

where $SP_i$ is the counterparty survival probability up to $t_i$. Assuming that the default is modelled by means of a generic intensity-based model, we can link survival probabilities to the so-called hazard rate function $h(t)$, (see e.g. [86]):

$$SP_i = \exp \left( - \int_{t_0}^{t_i} h(t) \, dt \right) .$$

A common assumption is to consider $h(t)$ constant between two consecutive dates in the set $\mathcal{T}$. The above approach, when computing CVA for a dealer (e.g. a bank), holds if there is no “market-credit” dependency. Instead it is possible that the probability of default of the counterparty increases with the exposure: in such a case one says that the dealer has a Wrong Way Risk (WWR) with its derivatives’ counterparty.

CVA is more important in relative terms when the derivative is uncollateralized. Here, we consider in detail WWR in such a situation and mention how to face the problem in the general case. A classical example is related to the exposure of a medium-sized corporate firm that has an amortizing loan with floating rate interest rate payments, whose interest rate risk has been managed writing an interest rate swap as payer counterparty with a dealer. Often such a loan has a horizon of 20 to 30 years and it constitutes a significant part of firm’s long term debt. If the swap has been closed before the 2007–2008 financial crisis, derivative’s Mark-to-Market $v(t)$ - a positive exposure for the dealer with current interest rates - could be of the same order of magnitude of the residual loan principal amount: this additional liability could increase corporate’s probability of default.

Another typical example among medium-sized corporates is described by an importer (exporter) which has hedged most of its foreign exchange exposure: a drop in sales and an adverse movement (w.r.t. the derivative) in the FX rate could endanger
financial stability. In both cases corporate probability of default could be significantly increased by a larger derivative’s value \( v(t) \).

In case of WWR a new, more sophisticated, model is needed because exposure and counterparty default probabilities are no more independent: exposure is conditional to default and a positive “market-credit” dependence originates the WWR. Recently an intuitive approach to WWR has been proposed in [67]: the conditional hazard rate is modelled as a stochastic quantity related to the portfolio value \( v(t) \) through a monotonic increasing function. In the following we focus on the specific functional form

\[
\hat{h}(t) = \exp \left( a(t) + b v(t) \right),
\]

where \( b \in \mathbb{R}^+ \) is the WWR parameter. In the following we call HW model this stochastic specification of hazard rate, even if results hold for a generic order-preserving function. The function \( a(t) \) is a deterministic function of time, chosen in such a way that on each date

\[
SP_i = \mathbb{E} \left[ \exp \left( - \int_{t_0}^{t_i} \hat{h}(t) \, dt \right) \right] \quad \forall i = 1, \ldots, n.
\]

(3.4)

Combining (3.3) and (3.4) we clearly see that function \( a(t) \) depends also on the value specified for the parameter \( b \).

The main advantage of HW model, besides the fact that is of easy understanding for non technical users and it is fast when running CVA computations, is that once one knows \( b \) and \( a(t) \), WWR can be implemented easily by a slight modification of existing algorithms for the calculation of CVA via the standard Basel approach (3.2). As already stressed in the literature, “an optimal model should leverage as much as possible from existing platform” for counterparty credit risk metrics (see [85]).

It is interesting to mention that \( v(t) \) in equation (3.3) does not need to be the derivatives’ portfolio value but it can also be related to the main market risk factors of the counterparty of interest: this fact, already mentioned in [67], has been discussed in detail in [85].

Unfortunately, in both cases the main difficulty in the proposed approach persists: given a WWR parameter \( b \), it is not straightforward to determine the function \( a(t) \). Let us stress that equation (3.4) is not elementary from a computational point of view since the sum is path dependent. This relevant point has already been mentioned by [67] when introducing the approach. In Section 3.3 we show a recursive relation that allows a simple determination of \( a(t) \) avoiding the path dependent nature of the quantities involved in the computation.

### 3.3 Model description

Let us assume that the market risk factors underlying the portfolio are discrete and we indicate with \( j_i \) the discrete state variable that describes the market at time \( t_i \). In this
Chapter 3. On the computation of CVA with WWR within the Hull-White Model

framework market dynamics is described by a Markov chain with

\[ q_i(j_{i-1}, j_i) \quad \forall i = 1, \ldots, n \]

the transition probability between \( j_{i-1} \) at time \( t_{i-1} \) and \( j_i \) at time \( t_i \). Typical examples where such a discrete approach is natural are lattice models like Cox-Ross-Rubinstein binomial method (CRR) [27] or the tree approach to the extended Vasicek model for interest rate derivatives [66]: both cases are analysed in detail in the next sections.

A discretization of equation (3.1), slightly simpler than (3.2), is

\[ CVA = (1 - R) \sum_{i=1}^{n} B_i EE_i PD_i \]

see e.g. equation (5) in [67]. For notational simplicity, except where differently stated, in the remaining part of the note we focus on this simpler approach [4]. The generalization of above CVA in presence of WWR is

\[ CVA_W = (1 - R) \sum_{i=1}^{n} E \left[ D_i E_i \tilde{PD}_i \right] \]

(3.5)

where \( \tilde{PD}_i \) is the stochastic probability to default between \( t_{i-1} \) and \( t_i \). Here we want to stress that expectation in (3.5) can be computed via any feasible numerical method: this fact implies that, given \( b \) and \( a(t) \), taking into account WWR just requires a slight modification in the payoff of existing algorithms used for the calculation of CVA.

We now introduce an innovative recursive approach that avoids the path dependency in the determination of \( a(t) \) so that equation (4) is satisfied. Hereinafter we refer to the technique to get such a function as either the calibration of \( a(t) \) or the "calibration problem": once the three sets of parameters (the recovery \( R \), the default probabilities \( PDs \) and the WWR parameter \( b \)) are estimated, this is the most complicated issue in the calibration of Hull-White model.

According to (3.3), the stochastic survival probability between \( t_{i-1} \) and \( t_i \) in the Hull-White model becomes

\[ \tilde{P}_i(j) \equiv \exp \left( -(t_i - t_{i-1}) \tilde{h}_i(j) \right) \equiv P_i \eta_i(j) \quad \forall i = 1, \ldots, n \quad , \]

(3.6)

where

\[ P_i \equiv \frac{SP_i}{SP_{i-1}} \]

is the forward survival probability between \( t_{i-1} \) and \( t_i \) valued in \( t_0 \). For notational convenience, we also set \( \tilde{P}_0(j_0) = \eta_0(j_0) = 1 \). The process \( \eta \) introduced in (3.6) can

\[ \text{Results can be easily generalized to the standard Basel approach (3.2).} \]
be seen as the driver of the stochasticity in survival probabilities and it plays a key role in circumventing path-dependency in the calibration of \( a(t) \), as shown in the following Proposition.

**Proposition 9** (Recursive approach to the calibration of \( a(t) \)). *In the model with discrete market risk factors, the calibration problem (3.4) becomes*

\[
\sum_{j_i} p_i(j_i) \eta_i(j_i) = 1 \quad \forall i = 1, \ldots, n ,
\]

where \( p_i(j_i) \) are probabilities and they can be obtained via the recursive equation

\[
p_i(j_i) = \sum_{j_{i-1}} q_i(j_{i-1}, j_i) \eta_{i-1}(j_{i-1}) p_{i-1}(j_{i-1}) \quad \forall i = 1, \ldots, n ,
\]

*with the initial condition* \( p_0(j_0 = 0) = 1 \).

*Proof.* See Appendix 3.A.

Thus the calibration problem (3.4) can be solved at each discrete date \( t_i \) via (3.7) by simply exploiting the fact that the process \( \eta \), non-path-dependent, is a martingale under the probability measure \( p \). Equation (3.8), in addition, specifies an algorithm to build this new probability measure recursively. In this framework \( \tilde{PD}_i \) can be readily obtained from (3.6). Let us mention that, although this is just one of the viable approaches to solve (3.4), it turns out to be, as shown in Section 3.5, a natural way to handle the additional complexity induced by early exercises within the Hull-White approach to WWR modelling.

### 3.4 Application to the case of linear derivatives

#### 3.4.1 The simplest application

A portfolio composed by only one FX forward derivative is the simplest example discussed in [67]. We consider a binomial tree, with all time intervals equal to \( \delta \). At each time step, there are two possible outcomes \( u = \exp(\sigma_x \delta^{1/2}) \) and \( d = 1/u \), where \( \sigma_x \) is the volatility of the underlying FX rate. The probability of one upward movement is

\[
q = \frac{1 - d}{u - d}
\]

and equal to \( 1 - q \) the probability of one downward movement. The discretization is very simple: log prices are regularly spaced and the transition probability is constant. It does depend neither from the time \( t_i \) nor from the positions \( j_{i-1} \) and \( j_i \), but only from the difference \( j_i - j_{i-1} \).
In this case the value of derivatives’ portfolio is
\[ v_i(j) = N F(0, t_n) (u^{2j-i} - 1) \]
where \( N \) is derivative’s notional and
\[ F(0, t_n) = S_0 \frac{B^f_n}{B^n} \]
is the FX forward rate at value date \( t_0 = 0 \), \( S_0 \) is FX spot rate, while the apex \( f \) stands for foreign.

This technique can be generalized to the case with collateral as in [67]. In this case the collateralized exposure is defined as
\[ E_C(t) = max[v(t) - C(t - \tau), 0] \]
where \( \tau \) is the cure period and \( C(t) \) the collateral posted at time \( t \). For example in the case considered in [67], where collateral is posted unilaterally by the counterparty, there is no rounding and no Minimum Transfer Amount, collateral assumes the following structure
\[ C(t) = max[v(t) - K, 0] \]
which takes into account a given positive threshold \( K \).

In order to compare the new technique with results in [67], we consider the same derivative contract and market conditions of [67]: a one year FX forward with \( N \) equal to 100 millions USD, a domestic and a foreign interest rate equal to 5\%, an unitary initial forward rate, a volatility \( \sigma_x \) equal to 15\%, \( b = 0.03 \) per million USD, a 125 bps flat counterparty credit spread \( c \) and a recovery \( R = 40\% \). As in [67], in this example we apply the approximation for the hazard rate \( h \approx c/(1 - R) \). For example in Table 3.1 we consider two different collateralization schemes with a zero threshold and a threshold \( K = 10 \) millions USD; collateral is posted unilaterally by the counterparty and cure period is 15 days. As shown in Table 3.1 results obtained are quite similar to values in [67]. In our computations we have chosen a daily lag, i.e. in the one year forward contract, given a value of \( b \), we obtain 365 piecewise constant values for \( a(t) \); unfortunately we do not know the number of discretization steps and the number of simulations used in examples of [67]. Even considering such a dense discretization, with the proposed approach computing time is negligible.

### 3.4.2 A relevant application

An interesting and relevant application is a derivatives’ portfolio with one receiver swap. For notational simplicity we consider the case of a swap with same frequency in fixed and floating legs, results can be extended to a generic interest rate swap. We
3.4. Application to the case of linear derivatives

<table>
<thead>
<tr>
<th></th>
<th>Long</th>
<th>No Collateral</th>
<th>K = 10</th>
<th>K = 0</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>CVA (mln USD)</strong></td>
<td>no path dependency</td>
<td>0.048</td>
<td>0.036</td>
<td>0.010</td>
</tr>
<tr>
<td>Values in [67]</td>
<td>0.048</td>
<td>0.036</td>
<td>0.011</td>
<td></td>
</tr>
<tr>
<td><strong>CV A_W/CVA − 1</strong></td>
<td>no path dependency</td>
<td>56.2%</td>
<td>41.7%</td>
<td>36.8%</td>
</tr>
<tr>
<td>Values in [67]</td>
<td>54.8%</td>
<td>41.7%</td>
<td>37.3%</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Short</th>
<th>No Collateral</th>
<th>K = 10</th>
<th>K = 0</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>CVA (mln USD)</strong></td>
<td>no path dependency</td>
<td>0.048</td>
<td>0.038</td>
<td>0.010</td>
</tr>
<tr>
<td>Values in [67]</td>
<td>0.048</td>
<td>0.039</td>
<td>0.011</td>
<td></td>
</tr>
<tr>
<td><strong>CV A_W/CVA − 1</strong></td>
<td>no path dependency</td>
<td>41.3%</td>
<td>34.3%</td>
<td>27.2%</td>
</tr>
<tr>
<td>Values in [67]</td>
<td>40.5%</td>
<td>34.0%</td>
<td>27.6%</td>
<td></td>
</tr>
</tbody>
</table>

Table 3.1: CVA in absence of WWR and the impact of the WWR on CVA with \(b = 0.03\) per million USD for both a long and a short forward position with maturity in one year. We compare the results obtained with our approach (no path dependency), with the ones in [67]. We consider both cases in absence and in presence of collateral with different thresholds (\(K\) equal either to 0 or to 10 millions USD), a cure period equal to 15 days and a notional equal to 100 millions USD, as in [67]. We have chosen a number of time intervals in the time discretization of the tree equal to \(n = 365\).

indicate with \(\{\hat{t}_i\}_{i=1,\ldots,v}\) swap reset dates and with \(s\) its fixed rate.

We consider a simple single-curve interest rate model as an extended Vasicek, where interest rate is modeled for its stochastic part by an Ornstein-Uhlenbeck process \(x(t)\) with a null long term mean (see e.g. [13])

\[
dx(t) = -\hat{a} x(t) \, dt + \hat{\sigma} \, dW(t)
\]  

(3.9)

where \(\hat{a}\) and \(\hat{\sigma}\) are the two main model parameters.

Discount factor at a future date \(t \in (0, T)\) is affine w.r.t. \(x(t)\), i.e.

\[
B(t, \hat{t}_i) = B(0; t, \hat{t}_i) \exp \left\{ -x(t) \frac{\sigma(t, \hat{t}_i)}{\hat{\sigma}} - \frac{1}{2} \int_{t_0}^{t} \left[ \sigma(u, \hat{t}_i) - \sigma(u, t) \right]^2 \, du \right\} \quad \forall \hat{t}_i \in (t, T)
\]

where \(B(0; t, \hat{t}_i)\) is the forward discount factor in \(t_0\) between \(t\) and \(\hat{t}_i\), while the stochastic discounts are

\[
D(t, \hat{t}_i) = B(t, \hat{t}_i) \exp \left\{ -\frac{1}{2} \int_{t}^{\hat{t}_i} \sigma(u, \hat{t}_i)^2 \, du - \int_{t}^{\hat{t}_i} \sigma(u, \hat{t}_i) \, dW(u) \right\}
\]

with

\[
\sigma(t, \hat{t}_i) = \frac{\hat{\sigma}}{\hat{a}} \left( 1 - e^{-\hat{a} (\hat{t}_i - t)} \right) \quad \forall \hat{t}_i > t
\]

Equation (3.9) can be accurately discretized by a trinomial tree (see [66]). As in previous subsection, time discretization intervals are all equal to \(\delta\), furthermore the grid is equally spaced by \(\Delta x = \sqrt{3} \delta \hat{\sigma}\). Transition probabilities do not depend on time \(i\)

\[
q_i(j_{i-1}, j_i) = q(j_{i-1}, j_i)
\]

101
and they assume values different from zero for a bounded set of $j_{i-1}$ and for just 3 values of $j_i - j_{i-1}$ (generally $0, \pm 1$), as described in [66]. In this case the transition probabilities are still very simple but more complicated than the ones described in Section 3.4.1. The value of the portfolio at time $t_i$ is

$$v_i(j) = N \left\{ \sum_{l=\kappa}^{\nu} (\hat{t}_l - \hat{t}_{l-1}) B(t_i, \hat{t}_l) - (1 + (\hat{t}_\kappa - \hat{t}_{\kappa-1}) L_f(\hat{t}_{\kappa-1}, \hat{t}_\kappa)) B(t_i, \hat{t}_\kappa) + B(t_i, t_n) \right\}$$

where $\hat{t}_\kappa$ is the first swap’s reset date after $t_i$, $N$ is swap notional and $L_f$ the Libor rate fixing at (previous) reset date $\hat{t}_{\kappa-1}$ that is paid in $\hat{t}_\kappa$.

### 3.4.3 Dataset description and numerical results

One of the main advantages of the described technique is that it allows a simple and fast calibration to market data. In this section, first we describe in detail the market dataset observed on September 13, 2012 and we mention the calibration techniques that we have adopted; then we show numerical results in the computation of CVA in presence of WWR for a yearly FX and a 5 years receiver swap vs 6m Euribor.

Interest rate data are from Bloomberg, while FX and CDS data are provided by Thomson Reuters. Market interest rates are the same of [11], EURUSD FX rate is equal to 1.2989, EURUSD FX volatility has been obtained by the 1 year ATM option with $\sigma_x = 10.159\%$. The (single) 6 months Euro curve has been bootstrapped from 6 months Euribor fixing and yearly swap rates on 6 months Euribor rate. Similarly the USD curve is obtained from the 3 months Libor fixing and yearly swap rates on 3 months Libor rate. Extended Vasicek parameters (in the Euro market) $\hat{a}$ and $\hat{\sigma}$ in (3.9) are calibrated to the diagonal volatilities of ATM swaptions with the sum of expiry and tenor equal to 5 years (see e.g. [13]), that are reported in Table 3.2. In particular we minimize the squared difference between market and model prices, obtaining $\hat{a} = 0.8771$ and $\hat{\sigma} = 0.0338$.

<table>
<thead>
<tr>
<th>Swaption Volatility (%)</th>
<th>1y4y</th>
<th>2y3y</th>
<th>3y2y</th>
<th>4y1y</th>
</tr>
</thead>
<tbody>
<tr>
<td>54.5</td>
<td>48.3</td>
<td>46.4</td>
<td>46.6</td>
<td></td>
</tr>
</tbody>
</table>

Table 3.2: Diagonal ATM swaption volatilities in the Euro market on September 13, 2012

We have considered as dealer counterparty ENI S.p.A. Its CDS term-structure on the same date is reported in Table 3.3 while its recovery $R$ is set equal to 40%. Given liquid credit market instruments, we can estimate survival probabilities and corresponding piecewise constant hazard rates have been obtained via the standard bootstrap technique from CDS spreads (see e.g. [56]). We show numerical results for a EURUSD FX forward with maturity in one year and a 5 years interest rate receiver swap on the 6 months Euribor traded at par on September
3.5. The Impact of Early Exercise

As already anticipated in section 3.1, CVA when early exercise is allowed gives rise to additional features. In this section we want to highlight the differences in CVA figures when both European and American options are considered, implementing the tree-based procedure described in Proposition 9. It is well known that backward induction and dynamic programming applied on (recombining) trees are, probably, the

<table>
<thead>
<tr>
<th>Tenor (years)</th>
<th>0.5</th>
<th>1.0</th>
<th>2.0</th>
<th>3.0</th>
<th>4.0</th>
<th>5.0</th>
<th>7.0</th>
<th>10.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>CDS Spread (bps)</td>
<td>18.70</td>
<td>29.86</td>
<td>65.18</td>
<td>100.83</td>
<td>121.89</td>
<td>138.80</td>
<td>151.53</td>
<td>161.67</td>
</tr>
</tbody>
</table>

Table 3.3: ENI mid-market CDS spreads in the Euro market on September 13, 2012

Figure 3.1: Impact of market-credit dependency in CVA (\(CVA_W/CVA\)) for the 1 year EURUSD FX forward on September 13, 2012 between a dealer and ENI S.p.A.

13, 2012 by a European dealer with ENI S.p.A. as corporate counterparty. The receiver swap has semiannual payments and an Act/360 daycount on both legs. In Fig. 3.1, we show the impact of market-credit dependency - measured as the ratio \(CVA_W/CVA\) - according to HW model with different values of \(b\) in the range \([0, 0.05]\) per million USD; in Fig. 3.2, a similar plot is shown in the interest rate swap case. In this last case we have considered for simplicity a 5 years Euro receiver swap versus the 6 months Euribor rate with same frequency and day-count (Act/360) on both legs; parameter \(b\) is in the range \([0, 0.05]\) per million EURO. The function \(a(t)\) obtained in the swap example with \(b = 0.03\) per million EURO is shown in Fig. 3.3. Results show that market-credit dependency could have a not-negligible impact on CVA computation.

3.5 The Impact of Early Exercise
Chapter 3. On the computation of CVA with WWR within the Hull-White Model

Figure 3.2: Impact of market-credit dependency in CVA ($CVA_W/CVA$) for the 5 years receiver interest rate swap on September 13, 2012 between a dealer and ENI S.p.A.

Figure 3.3: Plot of $a$ as a function of $t$ (in years) for the 5 years receiver Euro interest rate swap with $b = 0.03$ per million EURO. Both legs have the same frequency and day-count, par swap rate is equal to 1.020%.

simplest and most intuitive tool to price derivatives with an early exercise as American options. For these options, indeed, Monte Carlo techniques turn out to be computationally intensive in case of CVA: the exercise date, after which the exposure falls to zero, depends on the path of the underlying asset and on the exercise strategy. In such
3.5. The Impact of Early Exercise

a case we are asked to describe two random times: the optimal exercise time and the counterparty default time.

3.5.1 The Pricing Problem

Since our goal is to study the effects of early exercise clauses on CVA, we focus on the case of a dealer that enters into a long position\(^5\) on American style derivatives with a defaultable counterparty. That is, the dealer is the holder of the option and she has the opportunity to choose the optimal exercise strategy in order to maximize the option value. In particular, following \([12]\), we would need to differentiate between two possible assumptions depending on the effects of counterparty defaultability on the exercise strategy. The option holder would or would not take into account the possibility of counterparty default when she chooses whether to exercise or not. In the former case the continuation value (the value of holding the option until the next exercise date) should be adjusted for the possibility of default. However following the actual practice in CVA computation, we assume that counterparty defaultability plays no role in defining the exercise strategy of the dealer. This means that the pricing problem (before any CVA consideration) is the classical one for American options in a default-free world.

Let us assume to have a tree for the evolution of market risk factors\(^6\) up to time \(T\). Hereinafter, without loss of generality, we can set a constant time step \(\Delta t\) and denote the time partition on the tree by means of an index \(i\) in \(T = \{t_i\}_{i=0,...,n}\) with \(t_i = i \, \Delta t\).

We further introduce an arbitrary set of \(m\) exercise dates \(E = \{e_k\}_{k=1,...,m}\) with \(E \subseteq T\) at which the holder can exercise her rights receiving a payoff \(\phi_k\) that could depend on the specific exercise date \(e_k\). In this setting we can deal indistinctly with European \((m = 1)\), Bermudan \((m \in \mathbb{N})\) and American options \((m \to \infty)\). The standard dynamic programming approach then allows us to compute the derivative value at each node of the tree:

\[
v_i(j_i) = \begin{cases} 
\phi_m(j_i) & \text{for } i \text{ s.t. } t_i = e_m = T, \\
\max(c_i(j_i), \phi_k(j_i)) & \text{for } i \text{ s.t. } t_i \in E \setminus \{e_m\}, \\
c_i(j_i) & \text{otherwise},
\end{cases}
\]  
(3.10)

with \(c_i\) the continuation value of the derivative defined as

\[
c_i(j_i) = B(i, i + 1; j_i) \sum_{j_{i+1}} q_i(j_i, j_{i+1}) v_{i+1}(j_{i+1}),
\]  
(3.11)

where the sum must be considered over all possible \(t_{i+1}\)-nodes connected to \(j_i\) at time \(t_i\) and \(B(i, i + 1; j_i)\) is the future discount factor that applies from \(t_i\) and \(t_{i+1}\) possibly

\(^5\)A short option position does not produce any potential CVA exposure.

\(^6\)If we describe the dynamics of the price of a corporate stock, we assume - for the sake of simplicity - that such entity is not subject to default risk.
depending on the state variable $j_i$ on the tree.

We describe in detail the simple 1-dimensional tree, however, extensions to the 2-factor case (as, for example, the G2++ model in [13] or the recent dual curve approach in [68]) are straightforward. Once the derivative value is computed for all nodes and the WWR parameter $b$ is specified,[7] we can calibrate the auxiliary function $a(t)$ in (3.3) by means of the recursive approach illustrated in Section 3.3. The advantages of such an approach are, in this case, twofold: we avoid path-dependency in the calibration of $a(t)$, as in any other possible application, and we deal with early exercises via (3.10) and (3.11) in a very intuitive way.

3.5.2 The Plain Vanilla Case

We now want to assess the impact of early termination on CVA in order to understand the potential differences that could arise between European and American options from a counterparty credit risk management perspective.

In the first test we study the case of a plain vanilla option: we assume that the dealer buys a call option from a defaultable counterparty. Counterparty default probabilities are described in terms of a CDS flat curve at 125 basis points as in [67]. More precisely, with a flat CDS curve we can approximate quite well the survival probability between $t_0$ and $t_i$ as

$$SP_i = \exp \left( -\frac{s_i t_i}{1 - R} \right),$$

where $s_i$ is the credit spread relative to maturity $t_i$ and $R$ the recovery rate, equal to 40%. We further assume that trades are fully uncollateralized.[8] The underlying asset is lognormally distributed and represented by means of a Cox-Ross-Rubinstein binomial tree. We can thus apply the dynamic programming approach described above to price options on the tree and calibrate the function $a(t)$ recursively via (3.7). This procedure turns out to be quite fast: the Matlab coded algorithm takes less than 0.1 second to run on a 3.06 Ghz desktop PC with 4 GB RAM when $n = m = 500$. Figure 3.4 shows CVA profile[9] for both European and American call options as function of WWR parameter $b$ and for different levels of cost of carry. From standard non arbitrage arguments, we indeed know that the optimality of early exercise for plain vanilla call options is related to the cost of carry (defined as the net cost of holding positions in the underlying asset). In the Black-Scholes framework, the option holder could optimally decide to early exercise the derivative if and only if the cost of carry is strictly lower than the

---

[7] We refer the interested reader to the original paper [67] for an heuristic approach to determine the parameter and to [85] for comprehensive numerical tests with market data.

[8] Here we are interested in analysing the full exposure profile as function of early exercise opportunities. On the other hand, more realistic collateralization schemes can be taken into account in a straightforward manner within the described framework.

[9] Once $b$ and $a(t)$ are determined we can use whatever numerical technique to compute (3.5). Here we simply implement a simulation-based scheme that uses the tree as discretization grid. The number of generated paths is $10^5$.
3.5. The Impact of Early Exercise

![CVA profiles for European and American options as function of WWR parameter $b$ for several levels of cost of carry (CoC). Parameters are $S_0 = 100$, $K = 100$, $\sigma = 25\%$, $r = 1\%$, $T = 1$, $n = m = 500$. Counterparty CDS curve flat at 125 basis points.](image)

Figure 3.4: CVA profiles for European and American options as function of WWR parameter $b$ for several levels of cost of carry (CoC). Parameters are $S_0 = 100$, $K = 100$, $\sigma = 25\%$, $r = 1\%$, $T = 1$, $n = m = 500$. Counterparty CDS curve flat at 125 basis points.

risk-free interest rate\textsuperscript{[10]} As shown in Figure 3.4, CVA profiles are significantly different for European and American options when early exercise can represent the optimal strategy (black and dark gray lines). In particular the impact of WWR is significantly less pronounced for American options compared to the corresponding European ones. On the other hand, when early exercise is no more optimal the two options are equivalent: light gray lines in Figure 3.4 are undistinguishable from each other. In addition, the upward shift in CVA exposures is due to the fact that an increase in cost of carry (e.g. a reduction in the dividend yield) is reflected entirely in an augmented drift of the underlying asset dynamics that makes, \textit{ceteris paribus}, the call option more valuable.

The effect of early exercise on exposure profiles is depicted in Figure 3.5 where a possible underlying asset path is displayed along with the optimal exercise boundary (reconstructed on the binomial tree) and the corresponding value of European and American options. Until the asset value remains within the continuation region (the area below the dashed line), the two options have a similar value with the only difference given by the early exercise premium embedded in the American style derivative. However, if the asset value reaches or crosses the exercise boundary the exposure due to the American option falls to zero while the European option remains alive until maturity. From the definition of CVA (3.1), we can see that early exercise, if optimal, reduces the exposure of the holder to the counterparty default by shortening the life of the option. The ef-

\textsuperscript{[10]}The classical example is an option written on a dividend paying stock. This frame includes also a call option on a commodity whose forward curve is in backwardation or on a currency pair for which the interest rate of the base currency is higher than the one of the reference currency.
Chapter 3. On the computation of CVA with WWR within the Hull-White Model

Figure 3.5: The effect of early exercise on exposures. Parameters are $S_0 = 100$, $K = 100$, $\sigma = 25\%$, $r = 1\%$, CoC = $-2\%$, $T = 1$, $n = m = 500$. Left hand scale: Asset path (black solid line) and optimal exercise boundary (dashed line). Right hand scale: European option (light gray line) and American option (dark gray line).

Effect is even more pronounced when we introduce the WWR: early redemption, indeed, would occur as soon as the portfolio value is large enough with the consequence to eliminate the exposure just when counterparty default probabilities become more relevant. It is possible, then, to identify in the early termination clause an important mechanism that limits CVA charges, particularly when market-credit dependency is non-negligible as shown in [51] in the case without WWR. Any change that makes early exercise more likely tends to enhance such a mechanism. We see this effect in Figure 3.6 where we display the difference in CVA between European and American options as function of WWR parameter and option moneyness. With a given underlying asset dynamics, potential early exercise date is closer for more in the money options: the right of the holder is more likely to be exercised sooner. This shortens the life of the option and reduces both CVA charge (with respect to European options) and WWR sensitivity (with respect to the corresponding European option and the American options with lower moneyness). In this section we have shown that WWR can play a very different role for European and American options. In our opinion, however, WWR should be analysed on a case by case basis in order to determine its magnitude and the adequate capital charge: a 40% increase in standard CVA could overestimate the losses for an American option that can be optimally exercised in a short period while could be reductive in cases where early termination is less likely.
3.5. The Impact of Early Exercise

![Graph showing the difference in CVA between European and American options as a function of WWR parameter b and moneyness. Parameters are $S_0 = 100$, $\sigma = 25\%$, $r = 1\%$, $CoC = -2\%$, $T = 1$, $n = m = 500$. Counterparty CDS curve flat at 125 basis points.]

Figure 3.6: Difference in CVA between European and American options as function of WWR parameter $b$ and moneyness. Parameters are $S_0 = 100$, $\sigma = 25\%$, $r = 1\%$, $CoC = -2\%$, $T = 1$, $n = m = 500$. Counterparty CDS curve flat at 125 basis points.

3.5.3 The Bermudan Swaption Case

Probably the most relevant case of long position on options with early exercise opportunities in the portfolios of financial institutions is represented by Bermudan swaptions. Such exotic derivatives are, indeed, used by corporate entities to enhance the financial structure related to the issue of callable bonds. Often, by selling a Bermudan receiver swaption to a dealer, the callable bond issuer can reduce its net borrowing cost. Usually the swaption is structured such that exercise dates match the callability schedule of the bond. Let $\hat{T}$ be the bond maturity date. The dealer has the right, at any exercise date $e_k \in E \setminus \{e_m\}$, to enter into an interest rate swap with maturity $\hat{T}$, where she receives the fixed rate $K$ (equal to the fixed coupon rate of the bond) and pays the floating rate to the bond issuer with first payment made on date $e_{k+1}$. In our test we use the Euro interbank market data as of 13th September 2012 as given in [11]. We assume that the dealer buys a 10 year Bermudan receiver swaption where the underlying swap has, for simplicity, both fixed and floating legs with semiannual payments. The swaption can be exercised semiannually and its notional amount is Eur 100 million. We describe interest rates dynamics with a 1-factor Extended Vasicek model on a trinomial tree as illustrated in Section 3.4.2. Model parameters are calibrated to market prices of European ATM swaptions with overall contract maturity equal to 10 years as shown in Table 3.4.

As done for the plain vanilla option, we value the Bermudan swaption on the tree via

---

11Often the bond can be called at any coupon payment date after an initial lockout period.
Chapter 3. On the computation of CVA with WWR within the Hull-White Model

<table>
<thead>
<tr>
<th>Swaption</th>
<th>1y9y</th>
<th>2y8y</th>
<th>3y7y</th>
<th>4y6y</th>
<th>5y5y</th>
<th>7y3y</th>
</tr>
</thead>
<tbody>
<tr>
<td>Volatility%</td>
<td>40.4</td>
<td>37.6</td>
<td>35.1</td>
<td>32.8</td>
<td>30.8</td>
<td>27.7</td>
</tr>
</tbody>
</table>

Table 3.4: Diagonal implied volatility of European ATM swaptions used to calibrate the 1-factor Hull-White model. Calibrated parameters are $\hat{a} = 0.0146$ and $\hat{\sigma} = 0.0089$.

dynamic programming and calibrate the WWR model function $a(t)$. Once again the combined approach on the tree allows to perform both tasks in a negligible amount of time. Figure 3.7 reports the WWR impact for uncollateralized transactions struck at different levels of moneyness: at the money (swaption strike set equal to the market 10 years spot swap rate) and ±50 basis points. The upper graph reports the case with no initial lockout period while in the lower one we assume that the option cannot be exercised in the first 2 years. When the option can be exercised with no restrictions, we observe a moderate inverse relationship between moneyness and WWR impact due to the protection mechanism: the opportunity to early exercise when the exposure is large limits the effect of increased counterparty default probabilities. On the other hand, the introduction of a lockout period intensifies the WWR impact. Intuitively, by expanding the lockout period we move towards the limiting case of a European option. In this case the moneyness-WWR effect is reversed: the more in the money the option is, the more relevant the WWR effect becomes. During the lockout period the in-the-money option has a considerably higher exposure to counterparty default that cannot be mitigated via early termination.

3.6 Concluding Remarks

Nowadays WWR is a crucial concern in OTC derivatives transactions. This is particularly true for uncollateralized trades that a financial institution could have in place with medium-sized corporate clients. In the first part of the Chapter we have proposed a simple and effective method to take into account path dependency when calibrating the time dependent normalization function $a(t)$ of the model presented in [67]. We have introduced a simple change of measure and shown that in the new probabilities’ set the problem is not anymore path dependent: this allows a very fast computation of $a(t)$ component given the WWR parameter $b$. This new probability set can be computed via a simple recursion formula, following the approach described in equation (3.8). We have shown that some results (even in presence of collateral) described in [67] can be easily reproduced and we have shown in detail how to calibrate the model on real market data for two interesting examples (an FX forward and an IR swap). In our opinion the proposed approach could turn out to be a valuable tool to build a solid CVA modelling framework when the inherent dependency between market variables

\[ ^{12} \text{As above, we define it to be the ratio } CVA_{W} / CVA \text{ as given, respectively, by } (3.5) \text{ and } (1.2). \]
3.6. Concluding Remarks

and counterparty credit quality cannot be neglected. The approach is extremely powerful since it allows to use the existing platforms for CVA computations. Nowadays each dealer’s counterparty is characterized by two parameters - PDs and LGDs \((1 - R)\) in the notation of this note) - that allow to compute CVA in absence of Wrong Way Risk; just associating to each counterparty a third quantity \((b\) parameter in HW model) it is possible to generalize CVA computation to the case in presence of WWR using the same technology.

Figure 3.7: Impact of WWR on Bermudan receiver swaptions as function of WWR parameter \(b\) for several levels of moneyness. Upper panel: No lockout. Lower panel: 2-year lockout. Market data as of 13th September 2012. Counterparty CDS curve flat at 125 basis points.
Chapter 3. On the computation of CVA with WWR within the Hull-White Model

The presence of early termination clauses in vulnerable derivatives portfolios makes the CVA computation even more tricky. In the second part of the Chapter we have shown a simple and effective approach to deal with calibration and pricing of CVA within the Hull-White framework [67] for American or Bermudan options by extending the new calibration procedure to the dynamic programming algorithm required to take into account the free boundary problem inherent in the pricing of such derivatives. Numerical tests carried out underline the importance of adequate procedures to differentiate CVA profiles for European and American options. The possibility of early exercise, indeed, plays a remarkable role in mitigating the WWR: an undifferentiated CVA pricing for contingent claims with different exercise styles would then lead to severe misspecification of regulatory capital charges.

An interesting topic for further research would consider the impact of counterparty defaultability in defining the dealer’s optimal exercise strategy. Even if intuitive, this poses non-trivial problems mainly due to the interrelation among derivative pricing, WWR and calibration of function $a(t)$. It is our opinion, however, that the described framework could be extended in this direction.
Appendix to Chapter 3

3.A Proof of Proposition 9

In the given discrete model, the probability for a given \( i \)-steps path \( \{ j_m \}_{m=0,..,i} \) is

\[
Q_{\text{path}_i} = \prod_{m=1}^{i} q_{m}(j_{m-1}, j_{m})
\]

Let us define

\[
p_i(j_i) = \sum_{\{ j_m \}_{m=1,..,i-1}} \prod_{m=1}^{i} \left[ q_{m}(j_{m-1}, j_{m}) \eta_{m-1}(j_{m-1}) \right]
\]

this quantity can be written as

\[
p_i(j_i) = \sum_{j_{i-1}} q_{i}(j_{i-1}, j_i) \eta_{i-1}(j_{i-1}) \sum_{\{ j_m \}_{m=1,..,i-2}} \prod_{m=1}^{i-1} \left[ q_{m}(j_{m-1}, j_{m}) \eta_{m-1}(j_{m-1}) \right]
\]

and we obtain equation (3.8).

Equation (3.4) is equivalent to impose that

\[
1 = \sum_{\{ \text{path}_i \}} Q_{\text{path}_i} \prod_{m=1}^{i} \eta_{m}(j_{m}) = \sum_{\{ j_m \}_{m=1,..,i}} \prod_{m=1}^{i} q_{m}(j_{m-1}, j_{m}) \eta_{m}(j_{m})
\]

From above equation we get

\[
1 = \sum_{j_i} \eta_{i}(j_i) \sum_{\{ j_m \}_{m=1,..,i-1}} \prod_{m=1}^{i} \left[ q_{m}(j_{m-1}, j_{m}) \eta_{m-1}(j_{m-1}) \right]
\]

and then, using \( p_i(j_i) \) definition, we prove equation (3.7). Finally, we observe that quantities \( \{ p_i(j_i) \} \) are probabilities, since they are positive by definition being sum and product of positive quantities, and they sum up to 1

\[
\sum_{j_i} p_i(j_i) = \sum_{j_{i-1}} \eta_{i-1}(j_{i-1}) p_{i-1}(j_{i-1}) \sum_{j_i} q_{i}(j_{i-1}, j_i) = \sum_{j_{i-1}} \eta_{i-1}(j_{i-1}) p_{i-1}(j_{i-1}) = 1
\]

where we have used equation (3.8) and we have observed that \( \sum_{j_i} q_{i}(j_{i-1}, j_i) = 1 \forall j_{i-1} \), due to the definition of transition probability ♦
Part III

On the impact of relative performance on fund’s manager investment strategies
4.1 Introduction

The performance of a fund with respect to a benchmark (relative performance) affects the asset management through two different channels: the remuneration scheme of the manager, and the flow of funds which is typically a convex increasing function of the relative performance of the fund. In this paper we investigate how various types of relative performance remuneration schemes affect the asset management of the fund. We contribute to the debate on asset managers’ remuneration which is centered on two main issues: asset-based vs. performance-based remuneration, asymmetric vs. fulcrum type fees.

The debate on the effects of remuneration schemes on asset management is large with non conclusive results. In our analysis we concentrate our attention on a partial equilibrium analysis (we take the remuneration contract as given) and in particular on the incentives for the manager to take risk in excess with respect to the benchmark (large tracking error). For a general equilibrium/principal-agent analysis of the remuneration of the asset manager we refer among the others to [2, 36, 37, 79, 82]. We

1This Chapter is based on the paper On relative performance, remuneration and risk taking of asset managers jointly written with Emilio Barucci and Daniele Marazzina (Politecnico di Milano).
provide conditions for unbounded excess risk when the relative performance tends to zero (gambling) and we characterize the investment strategy with respect to the relative performance scheme. Our results suggest that the key point to prevent gambling is not the shape of a performance-based remuneration scheme (asymmetric vs. fulcrum fee) but rather the fact that manager is remunerated through a stake, a management/performance fee which is proportional to the assets under management (AUM)/relative performance with no fixed fee.

The regulation of the mutual fund industry in US constrains the remuneration scheme of asset managers to be of fulcrum type (centered around an index with increases in fees for a performance above the index matched by decreases in fees for a performance below the index) under the assumption that an asymmetric convex remuneration scheme (call option on the performance upside) would induce the manager to take risk in excess. In a partial equilibrium setting, the claim was theoretically confirmed by two classical contributions. In a lognormal environment, [58] show that a risk neutral asset manager remunerated through an asymmetric fee adopts a buy&hold strategy that maximizes the tracking error, instead a well designed fulcrum fee induces the manager to reduce the tracking error. [20] extends the analysis to the case of a risk averse manager remunerated through a call option with a constant strike, she shows that the investment strategy depends on the performance: the manager tends to take unbounded risk in excess with respect to the benchmark when the performance deteriorates significantly whereas the tracking error is limited when the performance is above the strike.

A second generation of contributions provides mixed evidence. Considering a remuneration scheme that is a nonlinear function of the relative performance, [8, 9, 22] show that a remuneration scheme with downside risk (non bounded from below pay, liquidation risk, personal capital investment) may induce the asset manager to take less risk (smaller tracking error) when his/her relative performance tends to zero with risk shifting incentives in a finite range (when the performance of the fund is below but not too far away from the benchmark). A robust result shows that the strategy converges to the Merton solution in case of very poor performance, whereas we observe convergence to a combination between the Merton solution and the benchmark when the performance is well above the benchmark. The result has been confirmed by [17, 35, 64] assuming an absolute performance remuneration scheme. Also the empirical evidence is mixed. [15] provide evidence that funds with poor performance are also characterized by higher variance of returns. [8, 18] find no evidence. [80] show that convexity induces the manager to take more risk.

In the above papers, the remuneration is a function of the relative performance either through the flow of funds or the remuneration scheme. In the first case we have an asset-based fee: [8] assume that the manager is remunerated through a management fee proportional to the AUM that are obtained by applying a multiplicative coefficient to the performance of the fund by the end of the horizon, the coefficient being a non-
linear function of the relative performance of the fund with respect to the benchmark. Building on the empirical literature, see [23], two types of flow fund functions have been analyzed: a symmetric (collar type) and an asymmetric one (call option type). [22] consider that the AUM are a smooth concave increasing function of the relative performance of the fund. In a different perspective, [31] and [72] consider the asset management problem assuming a performance fee that is a convex increasing function of the relative performance of the fund.

In our paper we consider remuneration schemes that include management/performance fees that are functions of the relative performance of the fund. In the first case, the management fee multiplied by the AUM is a nonlinear function of the relative performance, in the latter case the performance fee is a function of the relative performance. We analyze various types of schemes: linear, convex, collar, convex-concave, concave-convex. We build upon a claim of [8]: considering a particular class of remuneration scheme and comparing their results with those of [20], they claim that excess risk taking over a finite range occurs when the manager is always penalized for performance deteriorations with no fixed safety net independent of the fund value. We show two main results. First of all, considering a large class of remuneration schemes we qualify the above claim showing that the manager takes limited risk in excess when the performance of the fund tends to zero in case there is no fixed fee and the remuneration is sensitive to the performance in a neighborhood of the origin. Second, differently from what is claimed by the regulation of mutual funds and by [58], there is no significant difference between a convex and a collar type remuneration scheme when the performance is poor, the difference is observed in case of a positive performance: in the first case (convex type) the optimal strategy converges to the Merton solution, in the second case (collar type) to the benchmark. Playing with the convexity and the concavity of the remuneration scheme, we show different shapes of the portfolio strategy, e.g. it can be either decreasing in the relative performance or decreasing, increasing and then decreasing. We fully characterize the strategies.

Our analysis shows that unbounded (excess) risk taking when the relative performance tends to zero does not occur if two conditions are met: the asset manager is not endowed with a fixed fee and his/her fee is sensitive to a very poor relative performance. This result agrees with [84], where he observes “to make agents more willing to take risks there should be more of a focus on offering downside protection than on offering them upside potential”. This entails that thanks to the flow of funds, to prevent the manager to take risk in excess, a remuneration scheme based on a management fee proportional to the AUM or on a stake of the fund performs better than a fixed fee and an incentive (symmetric or asymmetric) fee based on the relative performance. Note that these results agree with empirical results obtained in [6], [53].

We provide an empirical analysis on the mutual funds industry. Remember that

\[ F(T)P(X(T)) \]
mutual funds are usually remunerated through a fixed fee and a management fee proportional to AUM, see [52]. Coherently with our analysis, and differently from [8], we show that there is a tendency to gambling with excess risk taking when the performance deteriorates significantly.

The paper is organized as follows. In Section 4.2 we introduce our setting. In Section 4.3 we analyze several different relative performance remuneration schemes, we derive the optimal portfolio solutions in closed form, and we compare the portfolio strategies. Finally, in Section 4.4 we present an empirical analysis on US equity mutual funds. All the proofs are postponed to the Appendix.

4.2 The model

We consider a continuous time economy. There are two assets: the risk-free asset with a constant instantaneous interest rate $r$ and a risky asset. The risky asset price $S(t)$ evolves as

$$dS(t) = S(t) (\mu dt + \sigma dZ(t)), \quad S(0) = S_0 > 0,$$  \tag{4.1}

where $\mu$ is the constant drift of the risky asset price, $\sigma$ is the constant volatility, $Z(t)$ is a one-dimensional Brownian motion on a complete probability space $(\Omega, \mathcal{F}, P)$. We denote by $\mathcal{F} = \mathcal{F}_t$ the $P$-augmentation of the filtration generated by $Z(t)$.

The manager is remunerated evaluating its performance against a benchmark $Y(t)$ which is a portfolio with a fraction $\beta$ invested in the stock market and $(1 - \beta)$ in the risk-free asset:

$$dY(t) = [(1 - \beta)r + \beta \mu]Y(t)dt + \beta \sigma Y(t)dZ(t), \quad Y(0) = Y_0 > 0.$$  

Given an adapted portfolio process $\theta(t)$, the performance of the the fund is

$$dF(t) = [(1 - \theta(t))r + \theta(t) \mu]F(t)dt + \theta(t) \sigma F(t)dZ(t), \quad F(0) = F_0 > 0;$$

we assume $F_0 = Y_0$. To have a well-defined problem, we require $\int_0^T (\theta(t) F(t))^2 dt < +\infty$ for any $T > 0$.

We consider the relative performance of the fund with respect to the benchmark $X(t) := \frac{F(t)}{Y(t)}$, which evolves as

$$dX(t) = X(t) \left[ \delta(t)(\mu - r - \sigma^2 \beta)dt + \delta(t) \sigma dZ(t) \right],$$  \tag{4.2}

with $\delta(t) = \theta(t) - \beta$ denoting the tracking error.

In the following we assume that the remuneration of the manager is defined over a finite horizon $T$ and that it depends on the terminal relative performance $X(T) := \frac{F(T)}{Y(T)}$. 


4.3. Relative Performance Remuneration Schemes

The manager defines the investment strategy $\theta(t)$ in order to maximize the expected utility of the remuneration. The utility function is of power type, i.e.

$$u(x) = \frac{x^a}{a}, \quad a < 1.$$  

Note that $1 - a$ represents the coefficient of relative risk aversion. The remuneration at time $T$ of the asset manager depends on the scheme which is a function of the relative performance at time $T$: $P(X(T))$.

4.3 Relative Performance Remuneration Schemes

We consider several types of remuneration schemes: linear, convex, collar, convex-concave, concave-convex. In detail we have the following schemes:

1. Linear: $P(X(T)) = K + mX(T), \; K \geq 0, \; m > 0$. The case $K = 0$ corresponds to a remuneration which linearly depends only on the relative performance of the fund with respect to the benchmark and $m$ represents the sensitivity of the remuneration to the extra performance with respect to the benchmark. $K$ represents the fixed fee.

2. Convex (or incentive, or call option): $P(X(T)) = K + (X(T) - 1)^+, \; K \geq 0$. The manager receives a fixed fee ($K$) and a variable component only in case the fund outperforms the benchmark at time $T$ $(X(T) \geq 1)$.

3. Collar type (or capped call option): $P(X(T)) = K + (X(T) - H)^+, \; K \geq 0, \; H > 1$. The manager receives a fixed fee $K$ and a variable component if the performance of the fund outperforms the benchmark $(X(T) > 1)$ with a cap at $X(T) = H > 1$ (the maximum remuneration is $K + H - 1$).

4. Convex-concave: $P(X(T)) = K + mX(T) + p(X(T) - 1)^+ - c(X(T) - H)^+, \; K, p, c \geq 0, \; m > 0, \; m + p - c > 0, \; H > 1$. The scheme is a combination of the linear and of the collar one. It is piecewise linear, the sensitivity of remuneration to the relative performance is low for a poor and for an outstanding performance compared to an intermediate performance $(1 \leq X(T) \leq H)$.

5. Concave-convex: $P(X(T)) = K + mX(T) + p(X(T) - H_1)^+ - c(H_2 - X(T))^+, \; K, p, c, m \geq 0, \; \min\{m + p, m + c\} > 0, \; H_1 > H_2 > 0, \; K - cH_2 \geq 0$. The scheme is specular to the Convex-concave. It is piecewise linear, the sensitivity of remuneration to the relative performance is high for a poor and for an outstanding performance compared to an intermediate performance.

The payoffs of the remuneration schemes are depicted in Figure 4.1. Note that the last two schemes are similar to those analyzed in [8, 9] but are not included in their framework.
Chapter 4. On relative performance, remuneration and risk taking of asset managers

Figure 4.1: Payoff schemes considered. Left: linear (blue dashed line), convex (red solid line) and collar (black dotted line). Right: concave-convex (red solid line) and convex-concave (dashed black line).

4.3.1 Linear scheme

The manager aims at solving the following problem:

$$\max_{\theta} E[u(K + mX(T))]$$

over the set of all admissible investment strategies $\theta$, subject to the dynamics for $X(t)$ given in (4.2). The above problem is related to the following value function

$$V(t, x) = \max_{\delta} E[u(K + mX(T))]|X(t) = x],$$

where, as above, $\delta$ is the tracking error with respect to the benchmark. This problem can be solved via the martingale technique, obtaining a closed form solution.

Theorem 4.3.1. Assuming $K > 0$, the optimal tracking error is

$$\delta^*(t) = \frac{\vartheta}{\sigma(1 - a)} \left[ 1 + \frac{K}{m} N(d_1(t)) \right],$$

where $X^*(t)$ denotes the optimal relative performance which evolves as

$$X^*(t) = \frac{K}{m} \left( \frac{N'(d_1(t))}{N'(d_2(t))} N(d_2(t)) - N(d_1(t)) \right),$$

with

$$\vartheta = \frac{\mu - r - \sigma^2 \beta}{\sigma},$$

and $N(\cdot)$ being the standard cumulative normal distribution (all the other coefficients are defined in the appendix).

If $K = 0$, we move to the classical Merton problem with terminal utility, and we obtain the constant optimal tracking error (Merton solution):

$$\delta^* = \frac{\mu - r - \beta \sigma^2}{(1 - a)\sigma^2},$$

(4.5)
4.3. Relative Performance Remuneration Schemes

Figure 4.2: Optimal tracking error $\delta^*$ as function of $a$ and $K$. Left: varying $a$ with $K = 1$. Right: varying $K$ with $a = 0.5$. Common parameters: $T = 2$; $t = 1$; $\mu - r = 0.03$; $\sigma = 0.2$; $\beta = 0.25$; $m = 1$.

i.e. Equation (4.3) with $K = 0$.

We observe that in case $\mu - r - \beta \sigma^2 > 0$ ($< 0$) the optimal investment strategy $\theta^* = \beta + \delta^*$ is overinvested (underinvested) in the risky asset with respect to the benchmark. In what follows we always assume parameter values such that $\mu - r - \beta \sigma^2 > 0$. Figure 4.2 (left) shows the optimal strategies varying $a$: the tracking error $\delta^*$ explodes as the manager’s risk tolerance increases ($a \to 1$) and it converges to zero in case risk aversion increases ($a \to -\infty$), the result agrees with the analysis in [72].

Figure 4.2 (right) shows the optimal strategy varying $K$. Note that a fixed fee $K > 0$ induces the manager to take excess (unbounded) risk when the relative performance of the fund deteriorates: the investment strategy is decreasing in $X$, it becomes unbounded as $X \to 0$, with the same sign as $\mu - r - \beta \sigma^2$, and the tracking error converges to the one obtained for $K = 0$ as $X \to \infty$. As $K$ increases, the tracking error increases, i.e. a higher floor induces the manager to take more risk.

4.3.2 Convex scheme

As in [58] and [20], we assume that the manager receives a fixed fee ($K$) and a variable component only in case the fund performance outperforms the benchmark at time $T$ ($X(T) \geq 1$). We model this component through a call option on $X(T)$ with strike price 1.

The asset manager aims at solving the following problem

$$\max_{\theta} E[u((X(T) - 1)^+ + K)],$$

subject to (4.2). Notice that $K \geq 0$, however, the additional condition $a \in (0, 1)$ is required if $K = 0$. The problem is similar to the one analyzed in [20] and the following result holds true.
Theorem 4.3.2. The optimal tracking error is
\[
\delta^*(t) = \frac{\vartheta}{\sigma X^*(t)} \left[ \frac{X^*(t)}{1 - a} + \frac{\vartheta N'(d_1(t))}{\sqrt{T - t}} - \frac{(1 - K) N(d_1(t))}{1 - a} \right],
\]
(4.6)
where
\[
X^*(t) = (1 - K) N(d_3(t)) + (\hat{x} - 1 + K) \frac{N'(d_1(t))}{N'(d_2(t))} N(d_2(t))
\]
(4.7)
(all the other coefficients are defined in the appendix).

The optimal strategy is depicted in Figures 4.3 and 4.4 and is similar to the one obtained in [20]. Notice that as \( X \to 0 \), \(|\delta^*|\) converges to infinity with a sign in agreement with the one of \( \mu - r - \beta \sigma^2 \). On the other hand, for large values of \( X \) the optimal strategy approaches the constant strategy obtained in case of a linear scheme with \( K = 0 \). The rationale of this shape is that for \( X \) large enough the remuneration is well approximated by the linear scheme (\( m = 1, K = 0 \)). Confirming what we have observed for the linear case, in Figure 4.3 we show that the risk exposure is higher the higher is the fixed fee \( K \). Risk exposure is also decreasing in risk aversion, see Figure 4.4.

4.3.3 Collar type scheme

We follow [58] considering a collar type remuneration scheme as the simplest fulcrum fee. The manager aims at solving the following problem
\[
\max_{\theta} E\left[u((X(T) - 1)^+ - (X(T) - \mathcal{H})^+ + K)\right], \quad \mathcal{H} > 1
\]
subject to (4.2). Without loss of generality, we assume a hurdle at 1. Again, the additional condition \( a \in (0, 1) \) is required if \( K = 0 \). The following result holds true.

Theorem 4.3.3. Assume \( \mathcal{H} > \hat{x} > 1 \), then, for any \( t \in (0, T] \) such that
\[
\max_{s \in [0,t]} X^*(s) < \mathcal{H}, \quad (4.8)
\]
the optimal tracking error is
\[
\delta^*(t) = \frac{\vartheta}{\sigma X^*(t)} \left[ \frac{X^*(t)}{1 - a} + \frac{\vartheta N'(d_3(t))}{\sqrt{T - t}} - \frac{(1 - K) N(d_3(t))}{1 - a} - (\mathcal{H} - 1 + K) \frac{N(d_1(t))}{1 - a} \right],
\]
where
\[
X^*(t) = (1 - K) N(d_3(t)) + (\hat{x} - 1 + K) \frac{N'(d_3(t))}{N'(d_4(t))} N(d_4(t)) + (\mathcal{H} - 1 + K) \left[ N(d_1(t)) - \frac{N'(d_1(t))}{N'(d_2(t))} N(d_2(t)) \right]
\]
(4.9)
(all coefficients are defined in the appendix).
4.3. Relative Performance Remuneration Schemes

Figure 4.3: Optimal tracking error $\delta^*$ for the linear, convex (option type) and capped (collar type) scheme. Parameters: $T = 2$; $t = 1$; $\mu - r = 0.03$; $\sigma = 0.2$; $a = 0.5$; $\beta = 0.25$. For the linear scheme we set $m = 1$, $K = 0$. For the other schemes we set $K = 0$ (left), $K = 0.2$ (right).

Figure 4.4: Optimal tracking error $\delta^*$ for the linear, convex (option type) and capped (collar type) scheme. Left: $a = 0.5$. Right: $a = -0.5$. Common parameters: $T = 2$; $t = 1$; $\mu - r = 0.03$; $\sigma = 0.2$; $\beta = 0.25$. For the linear scheme we set $m = 1$, $K = 0$. For the other schemes we set $K = 0.3$.

Condition (4.8) is due to the fact that if there exists a time instant $t \in [0, T]$ such that $X^*(t) = \mathcal{H}$, then the optimal strategy is trivial: for any $s \in [t, T]$, $X^*(s) = \mathcal{H}$, $\delta^*(s) = 0$.

In Figures 4.3 and 4.4 we plot the optimal tracking error $\delta^*$ considering the three remuneration schemes analyzed above. Note that the optimal tracking error for the convex and for the collar type remuneration scheme are similar in case of a poor performance (flat fee) and in the interval with a remuneration linear in the relative performance ($1 \leq X^*(t) \leq \mathcal{H}$). Instead, they look different when the cap is reached: the optimal strategy for a collar scheme exhibits a zero tracking error, the optimal strategy for a call option converges to the one obtained in case of a linear scheme. As for a convex scheme, in case of a collar type scheme risk exposure goes up with the safety net $K$ and as the risk aversion decreases.
Notice that although the relative performance remuneration scheme is exactly the one considered in [8], the optimal strategy looks quite different. In [8] the expected utility is a function of the AUM that are obtained by multiplying (according to our notation) \( P(X(T)) \) for \( F(T) \) (asset-based fee). In that framework, excess risk taking is observed over a finite range, there is no unlimited risk exposure in case of a very poor performance, and when the relative performance reaches the cap the optimal strategy tends to the Merton solution. The fact that the nonlinear relative performance function multiplies the funds (linear term) makes the difference driving the results: in [8] the remuneration is always sensitive to the fund performance, whereas in our setting it is flat in case of a poor and of an outstanding performance.

We can conclude that when the relative performance is considered per se with a fulcrum type remuneration scheme there is an incentive to gambling in a neighborhood of the origin. Our result is at odds with the analysis of [58] and provides a support to the regulation of the mutual fund industry.

### 4.3.4 Convex-concave scheme

We assume that the manager aims at solving the following problem

\[
\max \theta E[u((X(T) - 1)^+ - (X(T) - \mathcal{H})^+ + mX(T) + K)],
\]

subject to (4.2). Without loss of generality we have set \( p = c = 1 \). Note that remuneration is piecewise linear in the relative performance and the sensitivity of the remuneration is \( m \) for \( X(T) \leq 1 \) and for \( X(T) \geq \mathcal{H} \) and is \( m + 1 \) for \( 1 \leq X(T) \leq \mathcal{H} \), i.e. sensitivity is higher just above the benchmark. Moreover, the shape of the remuneration scheme is convex and then concave in the relative performance.

In this case, the concavification is not uniquely defined: two assumptions should be considered.

**Assumption 4.3.1.** There exists \( \hat{x}_u \in (1, \mathcal{H}] \) such that:

- \( U'(\hat{x}_u) = \frac{U(\hat{x}_u) - U(0)}{\hat{x}_u} \);
- \( U(x) \leq \frac{U(\hat{x}_u) - U(0)}{\hat{x}_u} x + U(0) \) for any \( x \in [0, \hat{x}_u] \);

being \( K > 0 \) if \( a \) is negative, i.e. \( U(0) \) is well-defined.

**Assumption 4.3.2.** There exist \( 0 < \hat{x}_d < 1 < \hat{x}_u \leq \mathcal{H} \) such that:

- \( U'(\hat{x}_u) = U'(\hat{x}_d) = \frac{U(\hat{x}_u) - U(\hat{x}_d)}{\hat{x}_u - \hat{x}_d} \);
- \( U(x) \leq \frac{U(\hat{x}_u) - U(\hat{x}_d)}{\hat{x}_u - \hat{x}_d} (x - \hat{x}_d) + U(\hat{x}_d) \) for any \( x \in [\hat{x}_d, \hat{x}_u] \).

The two assumptions are related to the two cases depicted in Figure 4.5.

The following results hold true.
4.3. Relative Performance Remuneration Schemes

Figure 4.5: Concavification for the convex-concave scheme. Left: \( m = 0.4 \). Right: \( m = 1 \). Common parameters: \( H = 5, a = K = 0.5 \).

Theorem 4.3.4. Let us assume that Assumption 4.3.1 holds true. Then the optimal tracking error is

\[
\delta^*(t) = \frac{\vartheta}{\sigma X^*(t)} \left( X^*(t) + \hat{x}_u \frac{N'(d_5(t))}{|\vartheta| \sqrt{T-t}} \right) - \frac{1-K N(d_5(t)) - N(d_3(t))}{1+m} - \frac{1-K \mathcal{H} N(d_1(t))}{1-a} \right),
\]

with

\[
X^*(t) = \frac{\hat{x}_u (1+m)}{1+a} \left( X^*(t) + \hat{x}_u \frac{N'(d_5(t))}{|\vartheta| \sqrt{T-t}} \right) - \frac{1-K N(d_5(t)) - N(d_3(t))}{1+m} \mathcal{H} \left( N(d_5(t)) - N(d_3(t)) \right) + \frac{1-K \mathcal{H} N(d_1(t))}{m} \left( N(d_5(t)) - N(d_3(t)) \right) + \frac{1-K \mathcal{H} N(d_1(t))}{m} \left( N(d_5(t)) - N(d_3(t)) \right) + \frac{1-K \mathcal{H} N(d_1(t))}{m} \left( N(d_5(t)) - N(d_3(t)) \right)
\]

(4.10)

(all coefficients are defined in the appendix).

Theorem 4.3.5. Let us assume that Assumption 4.3.2 holds true and \( K > 0 \). Then the optimal tracking error is

\[
\delta^*(t) = \frac{\vartheta}{\sigma X^*(t)} \left( X^*(t) + (\hat{x}_u - \hat{x}_d) \frac{N'(d_5(t))}{|\vartheta| \sqrt{T-t}} \right) + \frac{K (N(d_7(t)) - N(d_5(t)))}{m} \left( \mathcal{H} \left( N(d_3(t)) - N(d_1(t)) \right) - \frac{1-K \mathcal{H} N(d_1(t))}{m} \left( N(d_7(t)) - N(d_5(t)) \right) \right),
\]

with

\[
X^*(t) = \frac{K N'(d_7(t))}{m N'(d_8(t))} N(d_8(t)) - \frac{m \hat{x}_d + K N'(d_5(t))}{m} \left( \frac{N'(d_6(t))}{N(d_6(t))} \right) - \frac{K N'(d_7(t))}{m} \left( N(d_7(t)) - N(d_5(t)) \right)
\]

127
Remark 4.3.1. (all coefficients are defined in the appendix).

Figure 4.6: Optimal tracking error $\delta^*$ for the collar and convex-concave remuneration scheme. Left: $K = 0.3$. Right: $K = 0$. Common Parameters: $T = 2$; $t = 1$; $\mu - r=0.03$; $\sigma = 0.2$; $\beta = 0.25$; $\mathcal{H} = 4$; $a = 0.5$.

\begin{align*}
&+ \frac{\tilde{x}_a(1+m)+K-1}{N'(d_5(t))} \frac{N'(d_5(t))}{N(d_6(t))} - \frac{\mathcal{H}(1+m)+K-1}{N'(d_4(t))} \frac{N'(d_3(t))}{N(d_4(t))} \\
&+ \frac{1-K}{1+m} \frac{N(d_3(t))}{N(d_4(t))} + \mathcal{H} \frac{N(d_3(t)) - N(d_1(t))}{N(d_4(t))} \\
&+ \frac{\mathcal{H}(1+m)+K-1}{m} \frac{N'(d_3(t))}{N'(d_4(t))} \frac{N(d_2(t))}{N(d_4(t))} + \frac{1-K}{m} \frac{N(d_1(t))}{N(d_4(t))}
\end{align*}
(4.11)

(All coefficients are defined in the appendix).

Remark 4.3.1. If $K = 0$, Theorem 4.3.5 holds true with Equation (4.11) replaced by

\begin{align*}
X^*(t) &= \frac{1}{m} \left( \frac{\lambda \xi(t)}{m} \right)^{1/(\alpha-1)} e^{-a2^{(1-\alpha)^2}(T-t)} - \tilde{x}_a \frac{N'(d_5(t))}{N'(d_6(t))} N(d_6(t)) \\
&+ \frac{\tilde{x}_a(1+m)-1}{1+m} \frac{N'(d_5(t))}{N'(d_6(t))} N(d_6(t)) - \frac{\mathcal{H}(1+m)-1}{1+m} \frac{N'(d_3(t))}{N'(d_4(t))} N(d_4(t)) \\
&+ \frac{1}{1+m} \frac{N(d_3(t)) - N(d_1(t))}{N(d_4(t))} + \frac{\mathcal{H}(1+m)-1}{m} \frac{N'(d_1(t))}{N'(d_2(t))} \frac{N(d_2(t))}{N(d_4(t))} \\
&+ \frac{1}{m} \frac{N(d_1(t))}{N(d_4(t)).}
\end{align*}

Notice that the above equation and Equation (4.11) are related, since

\[ \lim_{K \to 0} K \frac{N'(d_7(t))}{N'(d_8(t))} N(d_8(t)) = \left( \frac{\lambda \xi(t)}{m} \right)^{1/(\alpha-1)} e^{-a2^{(1-\alpha)^2}(T-t)}. \]

In Figure 4.6 we plot the optimal tracking error for the collar and the convex-concave scheme (considering two different values of $m$ and of $K$). These figures highlight the role of a linear component ($m \neq 0$), which was absent in the convex and in the collar type scheme, and of a fixed floor ($K = 0$). We recall that the collar type fee corresponds
4.3. Relative Performance Remuneration Schemes

to the convex-concave case with \( m = 0 \). If \( K > 0 \) and \( m \neq 0 \), for all set of parameters (including \( p, c \neq 1 \)), then the optimal tracking error tends to become unbounded as \( X \to 0 \) yielding excessive risk taking. Instead, if \( K = 0 \) and \( m \neq 0 \), then the optimal tracking error converges to the Merton solution (4.5) as \( X \to 0 \). Note that the analysis of the collar scheme shows that the first shape is obtained also with \( K > 0 \) and \( m = 0 \). As wealth increases, the optimal strategies converge again to the optimal strategy in (4.5). The shape of the optimal strategy for \( K = 0 \) is the one of [8]. Notice that, as in [8], the optimal strategy exhibits a kink, due to the fact that the first order derivative of the concavified objective function is not well defined in \( \mathcal{H} \).

4.3.5 Concave-convex remuneration scheme

We consider a payoff similar to the one analyzed in [17] for an absolute performance fee: the manager aims at solving the following problem

\[
\max \theta E[u(p(X(T) - H_1)^+ + mX(T) - c(H_2 - X(T))^+ + K)],
\]

with \( H_1 > H_2 > 0 \) and \( K - cH_2 \geq 0 \). Again, in this case, the concavification is not uniquely defined. More precisely, two assumptions should be considered.

Assumption 4.3.3. There exists \( \hat{x}_u > H_1 \) such that:

\[
\bullet \ U'(\hat{x}_u) = \frac{U(\hat{x}_u) - U(0)}{\hat{x}_u};
\]

\[
\bullet \ U(x) \leq \frac{U(\hat{x}_u) - U(0)}{\hat{x}_u} x + U(0) \text{ for any } x \in [0, \hat{x}_u],
\]

being \( K - cH_2 > 0 \) if \( a \) is negative, i.e. \( U(0) \) is well-defined.

Assumption 4.3.4. There exist \( 0 < \hat{x}_d < H_2 < H_1 < \hat{x}_u \) such that:

\[
\bullet \ U'(\hat{x}_d) = U'(\hat{x}_d) = \frac{U(\hat{x}_d) - U(\hat{x}_u)}{\hat{x}_u - \hat{x}_d};
\]

\[
\bullet \ U(x) \leq \frac{U(\hat{x}_d) - U(\hat{x}_u)}{\hat{x}_u - \hat{x}_d} (x - \hat{x}_d) + U(\hat{x}_d) \text{ for any } x \in [\hat{x}_d, \hat{x}_u].
\]

The following results hold true.

Theorem 4.3.6. Let us assume that Assumption 4.3.3 holds true. Then the optimal tracking error is

\[
\delta^*(t) = \frac{\vartheta}{\sigma X^*(t)} \left( \frac{X^*(t)}{1 - a} + \hat{x}_u N'(d_1(t)) \frac{pH_1 - K N(d_1(t))}{m + p} \right),
\]

with

\[
X^*(t) = \frac{pH_1 - K}{m + p} N(d_1(t)) + \hat{x}_u (m + p) + K - pH_1 N'(d_1(t)) N(d_2(t)) \quad (4.12)
\]

(all coefficients are defined in the appendix).

As an example, Assumption 4.3.3 holds true if \( a = -0.5; \ m = 0.02; \ K = 0.5; \ c = 0.02; \ H_1 = 5; \ H_2 = 3; \ p = 0.2 \). If \( K = 0.2 \) Assumption 4.3.3 holds true (other parameters as above). Moreover, if \( K = 0.08 \) none of the two assumptions hold true: in this case a concavification is still possible in the sense of Assumption 4.3.4 with \( \hat{x}_d = H_2 \), and with a concavified utility function with a first order derivative not well defined in \( H_2 \) (see [17] for further details).
Chapter 4. On relative performance, remuneration and risk taking of asset managers

**Theorem 4.3.7.** Let us assume that Assumption 4.3.4 holds true and $K - cH_2 > 0$. Then the optimal tracking error is

$$\delta^*(t) = \frac{\vartheta}{\sigma X^*(t)} \left( \frac{N'(d_1(t))}{\sqrt{T-t}} \right)$$

$$+ \frac{X^*(t)}{1-a} \frac{pH_1 - K N(d_1(t))}{m+p} \frac{N'(d_1(t))}{N(d_2(t))} N(d_2(t)) + \frac{pH_1 - K N(d_1(t))}{m+p} N(d_1(t))$$

$$+ \frac{cH_2 - K}{m+c} \left( (K - cH_2) \frac{N'(d_1(t))}{N'(d_2(t))} N(d_2(t)) - ((m+c)\bar{x}_d + K - cH_2) \frac{N'(d_1(t))}{N'(d_2(t))} N(d_2(t)) \right)$$

with

$$X^*(t) = \frac{\bar{x}_u(m+p)+K - pH_1 N'(d_1(t))}{m+p} \frac{N'(d_2(t))}{N(d_2(t))} N(d_2(t)) + \frac{pH_1 - K N(d_1(t))}{m+p} N(d_1(t))$$

$$+ \frac{1}{m+c} \left( (K - cH_2) \frac{N'(d_3(t))}{N'(d_4(t))} N(d_4(t)) - ((m+c)\bar{x}_d + K - cH_2) \frac{N'(d_1(t))}{N'(d_2(t))} N(d_2(t)) \right)$$

$$+ \frac{cH_2 - K}{m+c} \left( N(d_3(t)) - N(d_1(t)) \right)$$

(4.13)

(All coefficients are defined in the appendix).

**Remark 4.3.2.** If $K = cH_2$, Theorem 4.3.4 holds true with Equation (4.13) replaced by

$$X^*(t) = \frac{\bar{x}_u(m+p) - pH_1 N'(d_1(t))}{m+p} \frac{N'(d_2(t))}{N(d_2(t))} N(d_2(t)) + \frac{pH_1}{m+p} N(d_1(t))$$

$$+ \frac{1}{m+c} \left( (\lambda\xi(t) \frac{1}{m+c} e^{\frac{a^2}{2N(1-a)^2}(t-t)} - ((m+c)\bar{x}_d - cH_2) \frac{N'(d_1(t))}{N'(d_2(t))} N(d_2(t)) \right)$$

$$+ \frac{cH_2}{m+c} \left( N(d_3(t)) - N(d_1(t)) \right).$$

We start by assuming $K - cH_2 > 0$. In Figure 4.7 we plot the optimal investment strategy varying $m$, $H_1$, $c$ and $K$. We would like to stress that Assumption 4.3.3 holds true only in the first figure (up-left), whereas Assumption 4.3.4 is satisfied in the other three cases. We observe that, as for the other remuneration schemes, the optimal tracking error converges to $\infty$ when the relative performance converges to zero, whereas the strategy converges to the optimal strategy obtained in the linear case for a relative performance high enough. In the transition, the shape can be either decreasing or decreasing-increasing-decreasing with a relative minimum and a relative maximum (hump shaped).

The non monotonic shape is due to the fact that the scheme is concave for a low relative performance. If this feature plays a relevant role, then the asset manager may take a limited risk exposure in case of a poor (but not extremely poor) performance. This interpretation is confirmed by the comparative statics analysis. If $m$ or $H_1$ are small enough, or $K$ is large enough, then the optimal strategy is decreasing as in the case of a convex scheme, otherwise the optimal strategy exhibits a local minimum.
4.4. Empirical Analysis

Figure 4.7: Optimal tracking error $\delta^*$ for the concave-convex remuneration scheme, varying $m$ (upper-left), $H_1$ (upper-right), $c$ (lower-left) and $K$ (lower-right). Common Parameters: $T = 2; t = 1; \mu - r = 0.03; \sigma = 0.2; \beta = 0.25; \alpha = 0.5; c = 0.02; H_1 = 4; H_2 = 1; p = 0.75; m = 0.3$ and $K = 0.2$.

and a local maximum. The interpretation is as follows. When the second kink at $H_1$ inducing convexity is high enough, we may observe a limited risk exposure for a poor performance (the concave part of the scheme plays a role). As far as the fixed fee $K$ is concerned, we confirm that a large fixed fee induces the manager to take a significant risk exposure and this effect is likely to induce a monotonic decreasing shape.

The shape of the optimal strategy changes significantly in case $K - cH_2 = 0$ (zero floor). As shown in Figure 4.8, if the fixed fee is null then the optimal strategy is characterized by excess risk taking over a finite interval; instead if there is a strictly positive floor to the remuneration then the asset manager takes an unlimited risk exposure as the relative performance tends to zero.

4.4 Empirical Analysis

The above theoretical analysis provides us with a research question that is worthwhile to be investigated: do relative performance remuneration schemes provide an incentive to take unbounded risk when the relative performance deteriorates ($H_p^0$), as suggested by [20], or a risk shifting incentive only over a finite range as suggested by [8] ($H_p^1$)?
Chapter 4. On relative performance, remuneration and risk taking of asset managers

Figure 4.8: Optimal tracking error $\delta^*$ for the concave-convex remuneration scheme. Left: $K = cH_2$. Right: $K = cH_2 + 0.1$. Common parameters: $T = 2$; $t = 1$; $\mu - r = 0.03$; $\sigma = 0.2$; $\beta = 0.25$; $a = 0.5$; $c = 0.02$; $H_1 = 4$; $H_2 = 1$; $p = 0.75$.

We address this question analyzing the performance of mutual funds replicating the analysis developed in [8]. We focus our analysis on US equity mutual funds and in particular we select the funds active in the period 1995 – 2014 more exposed to equity-based active management, i.e. those classified as Capital Appreciation, Growth and Income or Growth Funds within the Lipper Funds Classification. For each year we restrict our analysis to only those funds with beta higher than one. In [25], authors show that funds more susceptible to benchmarking pressures increase their demand for high beta stocks reducing their exposure to low-beta ones. Since benchmarking plays a crucial role in our model, then, this choice seems to be adequate for describing the framework our model can be as closely as possible applied to.

Using daily data we drop funds with less than 250 observations and funds no more active at the end of the period in exam. We use the S&P500 index as benchmark for all funds and the rate on 3 months T-bills as proxy of risk-free rate. All financial data are obtained from the Thomson Reuters database.

Note that in the mutual fund industry the remuneration scheme usually consists in a fixed fee and in a management fee proportional to AUM, see [52]. According to our analysis we should find evidence in favor of $H_p^0$. We want to analyze how different levels of relative performance affect the manager’s risk-shifting incentives.

We first focus on underperforming funds (Table 4.1). At a monthly level we split the range of underperformance with respect to the benchmark into mutually exclusive intervals delimited by the reference levels $B = \{0\%, 5\%, 10\%, 15\%, 20\%\}$. That is, we associate, for example, a 6.5% monthly underperformance to the 5%-interval. The last interval, i.e. the one bounded from above by $-20\%$, contains all values of monthly underperformances equal to or higher than 20%. For each interval we create a corresponding dummy variable $\text{UNDER}(b)_{i,m}$ which is equal to one if the fund’s relative performance in the preceding month is in the $b$-interval and zero otherwise.
4.4. Empirical Analysis

Under $Hp^0$ we expect that the farther a fund falls behind the benchmark, the higher the incentives that a manager has to take risk in excess detaining a larger stake in the risky asset with respect to the benchmark. As shown in Section 4.3, this effect is obtained for all the remuneration schemes, except in case there is no fixed fee and the remuneration is sensitive to the performance in a neighborhood of the origin (concave-convex or convex-concave remuneration schemes). In the latter case, risk exposure is likely to be hump shaped: limited exposure when the relative performance is significantly poor, significant exposure for intermediate poor relative performance, again limited exposure when the performance falls behind the benchmark but not too much.

To test the two hypotheses, we regress the daily fund’s excess returns $R_{F,t} - r_t$ on the S&P500 index’s daily excess returns $R_{B,t} - r_t$ interacted with the $UNDER$ indicators defined above. Performing this regression we implicitly assume that the benchmark of the funds is related to the S&P500 index. Here $r_t$ denotes the risk-free rate observed on day $t$. We also include year and month fixed effects to address the seasonality in betas, see for example [76]. Regression results are shown in Table 4.1. Estimated coefficients are all positive and increasing in the magnitude of underperformance: funds falling far below the market at the end of a month tend to take excess risk by increasing their investment in the risky asset with respect to funds with better relative performance. These evidences, combined with the fact that we restrict our analysis only to funds with beta higher than one, can be interpreted in favour of the unbounded risk taking incentive hypothesis $Hp^0$. Indeed, gambling incentives become more and more relevant as relative performance deteriorates (increasing coefficients in the magnitude of the underperformance), leading the underperforming managers (whose portfolio beta is already higher than one) to further deviate from the benchmark risk profile.

We complete our analysis looking at the investment strategy of overperforming funds (Table 4.2). The analysis developed in Section 4.3 is not fully conclusive. As a matter of fact, in case of a collar type remuneration scheme we have that the investment strategy is monotonic and converges to the benchmark. Instead, in case of the other remuneration schemes we have that the optimal investment strategy converges to the one adopted in case of a linear remuneration scheme. The strategy foresees an investment in the risky asset in excess with respect to the benchmark if and only if $\mu - r - \beta \sigma^2 > 0$, otherwise the manager invests in the risky asset less than the benchmark. In the first case we expect a limited exposure to the benchmark (decreasing in the overperformance of the fund), in the other cases the strategy should closely track the benchmark.

We have performed a regression analysis similar to the one developed for the case of underperforming funds. Analyzing the results exhibited in Table 4.2, we note that,

---

4For the sake of completeness, we point out that these results are not directly comparable to those obtained in [8] where the findings support the hypothesis $Hp^1$. [8], indeed, concentrate their analysis on funds with beta lower than one in order to isolate (sufficiently) risk averse managers.
Chapter 4. On relative performance, remuneration and risk taking of asset managers

### Table 4.1:
Regression of daily funds’ excess returns on S&P500 index’s daily excess returns interacted with UNDER indicators for the subset of funds whose current-year beta is higher than 1. The fixed effects variables are interacted with \((R_f^B - r_t)\). Numbers in parenthesis correspond to t-statistics computed with robust and clustered standard errors. The row Wald Test displays p-values for the null hypothesis that the coefficients in the corresponding regression are jointly equal.

<table>
<thead>
<tr>
<th>UNDER(0%)(_{i,m}) × ((R_f^B - r_t))</th>
<th>0.0071</th>
<th>0.0019</th>
<th>-0.0002</th>
<th>0.0008</th>
<th>0.0156</th>
</tr>
</thead>
<tbody>
<tr>
<td>(4.92)</td>
<td>(1.42)</td>
<td>(-0.16)</td>
<td>(-0.60)</td>
<td>(15.94)</td>
<td></td>
</tr>
<tr>
<td>UNDER(5%)(_{i,m}) × ((R_f^B - r_t))</td>
<td>0.0936</td>
<td>0.1096</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(21.74)</td>
<td>(24.20)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UNDER(10%)(_{i,m}) × ((R_f^B - r_t))</td>
<td>0.1334</td>
<td>0.1604</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(14.28)</td>
<td>(16.59)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UNDER(15%)(_{i,m}) × ((R_f^B - r_t))</td>
<td>0.1878</td>
<td>0.2206</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(6.70)</td>
<td>(7.78)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UNDER(20%)(_{i,m}) × ((R_f^B - r_t))</td>
<td>0.3626</td>
<td>0.3989</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(19.78)</td>
<td>(20.01)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| Month fixed effects | Yes | Yes | Yes | Yes | Yes |
| Year fixed effects | Yes | Yes | Yes | Yes | Yes |
| \(R^2\) | 0.84 | 0.84 | 0.84 | 0.84 | 0.84 |
| Wald Test | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| N. of obs | 4397622 |

for low to moderate underperformances, coefficients are positive and increasing with the level of relative performance. Then, as soon as overperformance is higher than 10%, coefficients decrease till becoming negative in the last row of Table 4.1, i.e. for outperformances that exceed 20%. These results agree with a collar type remuneration scheme and may suggest that the asset managers are capped in their remuneration yielding a lower exposure to the benchmark.

### 4.5 Conclusions

The debate on misalignment between the asset manager behavior and the investors’ interests is intense and non conclusive. The main issue concerns the incentives to take excessive risk by the manager which may be due to the composition of the remuneration scheme. The debate mainly regards two features: asset-based vs. performance-based remuneration schemes, asymmetric vs. fulcrum type fees.

There is a well established claim that asymmetric fees based on the performance induce the asset manager to take risk in excess whereas a fulcrum type fee should prevent it. On the other hand, asset-based fees should perform better than performance-based fees in preventing excessive risk taking but may not be efficient in aligning interests of asset managers and of investors.

Concentrating our attention on remuneration schemes related to the performance of the fund with respect to a benchmark, we have shown that it is not the asymmetric-
### 4.5. Conclusions

<table>
<thead>
<tr>
<th>Dependent Variable: $R_{i,t}^f - r_t$</th>
<th>( \text{OVER}(0%)<em>{i,m} \times (R^f</em>{i,t} - r_t) )</th>
<th>( \text{OVER}(5%)<em>{i,m} \times (R^f</em>{i,t} - r_t) )</th>
<th>( \text{OVER}(10%)<em>{i,m} \times (R^f</em>{i,t} - r_t) )</th>
<th>( \text{OVER}(15%)<em>{i,m} \times (R^f</em>{i,t} - r_t) )</th>
<th>( \text{OVER}(20%)<em>{i,m} \times (R^f</em>{i,t} - r_t) )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.0312 0.0326 0.0328 0.0329 0.0309</td>
<td>0.0539 0.0548</td>
<td>0.0847 0.0889</td>
<td>0.0717 0.0776</td>
<td>-0.1347 -0.1259</td>
</tr>
<tr>
<td></td>
<td>(26.68) (27.81) (27.46) (27.34) (27.74)</td>
<td>(7.77)</td>
<td>(3.04) (3.15)</td>
<td>(1.26) (1.25)</td>
<td>(-2.11) (-1.95)</td>
</tr>
</tbody>
</table>

Month fixed effects | Yes | Yes | Yes | Yes | Yes |
Year fixed effects | Yes | Yes | Yes | Yes | Yes |
\( R^2 \) | 0.84 | 0.84 | 0.84 | 0.84 | 0.84 |
Wald Test | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
N. of obs | 4397622 |

**Table 4.2:** Regression of daily funds’ excess returns on S&P500 index’s daily excess returns interacted with \( \text{OVER} \) indicators for the subset of funds whose current-year beta is higher than 1. The fixed effects variables are interacted with \( (R^f_{i,t} - r_t) \). Numbers in parenthesis correspond to t-statistics computed with robust and clustered standard errors. The row Wald Test displays p-values for the null hypothesis that the coefficients in the corresponding regression are jointly equal.

A fulcrum type feature that makes the difference in preventing excessive risk taking in case of a poor performance. To prevent gambling when the performance deteriorates, it is important that the asset manager is not endowed with a fixed fee and that his/her remuneration is sensitive to a very poor relative performance.
To prove the theorems, let us consider an economy with zero risk-free rate \( r^\ast = 0 \), a risky asset with drift \( \mu^\ast = \mu - r - \beta \sigma^2 \) and volatility \( \sigma \). Therefore the evolution of the relative performance process \( X(t) \) in (4.2) is fully described by

\[
dX(t) = X(t) \left[ \delta(t)(\mu^\ast dt + \delta(t)\sigma dZ(t)) \right]
\]

and thus in this new economy \( X \) is the process which describes the portfolio value (if the agent can invest in both the risky and the risk-free asset) and \( \delta \) is the amount of wealth (in percentage) invested in the risky asset.

In this framework the state price density is given by

\[
\xi(t) = e^{-(\vartheta + |\vartheta|^2/2)t - \vartheta Z(t)}
\]

with \( \vartheta = \frac{\mu - r^\ast}{\sigma} = \frac{\mu - r - \sigma^2 \beta}{\sigma} \).

The outline of all the proofs is the following, see [71] for details: once an extended concave utility function \( \tilde{U} \) is defined, we compute its set-valued first order derivative \( \tilde{U}' \) as well as the function \( I \), which is the inverse of \( \tilde{U}' \) in the sense that \( z \in \tilde{U}'(I(z)) \).

The starting problem is equivalent to the following:

\[
\max_{X(T)} E\left[ \tilde{U}(X(T)) \right] \quad \text{such that} \quad E[\xi(T)X(T)] \leq X(0) \quad \text{and} \quad X(T) \geq 0.
\]

Therefore the martingale approach considered, for example, in [20] can be extended to solve our problems, with \( X^\ast(T) = I(\lambda \xi(T)) \), and with \( \lambda \), the Lagrangian multiplier, solution of

\[
E[\xi(T)I(\lambda \xi(T))] = X(0) = 1.
\] (4.14)

We would like to stress that the function \( \Xi(\lambda) = E[\xi(T)I(\lambda \xi(T))] \) is continuous and strictly decreasing. Furthermore \( \lim_{\lambda \to +\infty} \Xi(\lambda) = 0 \) and \( \lim_{\lambda \to 0} \Xi(\lambda) = +\infty \). Therefore a solution of Equation (4.14) always exists.

The optimal process \( X^\ast \) is then computed for any \( t \in [0, T] \) as

\[
X^\ast(t) = E_t \left[ \frac{\xi(T)}{\xi(t)} X^\ast(T) \right] = E_t \left[ \frac{\xi(T)}{\xi(t)} I(\lambda \xi(T)) \right].
\] (4.15)

Notice that the process \( X^\ast \) is the optimal process also for the initial problem (with non-concave utility function \( U \)) due to [20 Equations (A5)-(A8)].
Appendix 4.A Proofs

To deal with the optimal allocation process, let us now define \( M(t) = \xi(t) X^*(t) \); being \( M \) a martingale \([71]\), there exists a function \( \varphi \) such that \( dM(t) = \varphi(t) dZ(t) \). Once \( \varphi \) is computed, due to \([71, \text{Theorem 3.7.3}]\), we have

\[
\delta^* = \frac{1}{\sigma} \left( \vartheta + \frac{\varphi(t)}{M(t)} \right). 
\]

In the following, we also denote with \( i \) is the inverse of the first order derivative of \( u \), i.e. \( i(z) = (u')^{-1}(z) = z^{1/(a-1)} \).

4.A.1 Proof of Theorem 4.3.1

Let us define the extended utility function \( \tilde{U}(x) = \begin{cases} U(x) := (mx + K)^a / a & \text{if } x \geq 0, \\ -\infty & \text{if } x < 0, \end{cases} \)

If \( m = 1 \) and \( K = 0 \) we obtain the classical Merton problem with terminal utility. The utility function \( \tilde{U} \) is not differentiable in 0, therefore we can define a set-valued function \( \tilde{U}' : [0, +\infty) \rightarrow (0, +\infty) \) by

\[
\tilde{U}'(x) = \begin{cases} \{U'(x)\} & \text{if } x > 0, \\ [U'(0), +\infty) & \text{if } x = 0, \end{cases}
\]

and its inverse

\[
I(z) = \frac{1}{m} \left( i \left( \frac{z}{m} \right) - K \right) 1_{z < U'(0)}.
\]

Equation (4.15) implies

\[
X^*(t) = -\frac{K}{m} E_t \left[ \frac{\xi(T)}{\xi(t)} 1_{\xi(T) < \gamma} \right] + \frac{1}{m} E_t \left[ \frac{\xi(T)}{\xi(t)} i \left( \frac{\lambda \xi(T)}{m} \right) 1_{\xi(T) < \gamma} \right]
\]

with \( \gamma = U'(0)/\lambda \), i.e. Equation (4.4), since

\[
E_t \left[ \frac{\xi(T)}{\xi(t)} 1_{\xi(T) < \gamma} \right] = N(d_1(t))
\]

with \( d_1(t) = (\ln(\gamma/\xi(t)) - |\vartheta|^2(T-t)/2) / (|\vartheta|\sqrt{T-t}) \), and

\[
E_t \left[ \frac{\xi(T)}{\xi(t)} i \left( \frac{\lambda \xi(T)}{m} \right) 1_{\xi(T) < \gamma} \right] = K \frac{N'(d_1(t))}{N'(d_2(t))} N(d_2(t)),
\]

with \( d_2(t) = d_1(t) + |\vartheta| \sqrt{T-t} / (1 - a) \).

Finally, computations lead to

\[
\varphi(t) = \vartheta \xi(t) \left( \frac{a}{1-a} X^*(t) + \frac{K}{m} \frac{N(d_1(t))}{1-a} \right),
\]

and therefore Equation (4.16) gives Equation (4.3).
4.A.2 Proof of Theorem 4.3.2

This result follows directly from [20]. More precisely, (4.6) and (4.7) correspond to [20, Equation (26)] and [20, Equation (25)], respectively. More precisely, in [20] the author solves the optimization problem

\[
\max_{X(T)} \mathbb{E}[u(\alpha(X(T) - B_0e^{rT})^+) + K]
\]

which corresponds to our problem (in the new economy) setting \( \alpha = 1 \), and \( B_0 = 1 \). Notice that in [20] \( K \) should be positive: however the results can be easily extended to the case \( K = 0 \), assuming \( a \in (0, 1) \).

More precisely, since the utility function \( U(x) := u((x - 1)^+ + K) \) is not concave in \( x \), we define the new concavified utility function

\[
\tilde{U}(x) = \begin{cases} 
U(x) & \text{if } x > \hat{x}, \\
U(0) + U'(\hat{x})x & \text{if } 0 \leq x \leq \hat{x}, \\
-\infty & \text{if } x < 0,
\end{cases}
\]

Let us consider the case \( K = 0 \): in order to have a continuous concavified utility function \( \tilde{U} \) it must be \( \hat{x} = 1/(1-a) \), which is greater than 1 since \( a \in (0, 1) \) (moreover, if \( K = 0 \) and \( a \) is negative, \( U(0) = -\infty \)). For the case \( K > 0 \) a similar result holds true with the existence (and uniqueness) of \( \hat{x} > 1 \) such that \( U(\hat{x}) = U(0) + U'(\hat{x})\hat{x} \), thanks to [20, Lemma 1].

The concavified utility function \( \tilde{U} \) is not differentiable in 0, therefore we can define a set-valued function \( \tilde{U}' : [0, +\infty) \rightarrow (0, +\infty) \) by

\[
\tilde{U}'(x) = \begin{cases}
\{U'(x)\} & \text{if } x > \hat{x}, \\
\{U'(\hat{x})\} & \text{if } 0 < x \leq \hat{x}, \\
[U'(\hat{x}), +\infty) & \text{if } x = 0.
\end{cases}
\]

See [20, Appendix]. We can also define the inverse for the function \( \tilde{U}' \) given by

\[
I(z) = (i(z) + 1 - K)1_{z < U'(\hat{x})}
\]

where \( I \) is necessary to compute \( \lambda \) exploiting Equation (4.14), and thus to get the coefficients in (4.6) and (4.7).

\[
d_1(t) = \left( \ln(\gamma/\xi(t)) - |\vartheta|^2(T - t)/2 \right) / \left( |\vartheta|\sqrt{T - t} \right),
\]

\[
d_2(t) = d_1(t) + |\vartheta|\sqrt{T - t}/(1 - a),
\]

being \( \gamma = u'(\hat{x} - 1 + K)/\lambda \).
4.A.3 Proof of Theorem 4.3.3

Since the utility function $U(x) := u((x - 1)^+ - (x - \mathcal{H})^+ + K)$ is not concave in $x$, we compute its concavification as in Equation (4.17). The concavified utility function $\tilde{U}$ is not differentiable in 0 and $\mathcal{H}$, therefore we can define a set-valued function $\tilde{U}': [0, +\infty) \rightarrow (0, +\infty)$ by

$$
\tilde{U}'(x) = \begin{cases} 
0, & U'(\mathcal{H}^-) \text{ if } x \geq \mathcal{H}, \\
U'(x) \text{ if } \hat{x} < x < \mathcal{H}, \\
U'(\hat{x}) \text{ if } 0 < x \leq \hat{x}, \\
[U'(\hat{x}), +\infty) \text{ if } x = 0,
\end{cases}
$$

where $\hat{x}$ is as in Appendix 4.A.2. We can also define an inverse for the function $\tilde{U}'$ given by

$$
I(z) = (i(z) + 1 - K)1_{U'(\mathcal{H}^-) \leq z < U'(\hat{x})} + \mathcal{H}1_{z < U'(\mathcal{H}^-)}.
$$

Let us define $\gamma_1 = U'(\mathcal{H}^-)/\lambda$, and $\gamma_2 = U'(\hat{x})/\lambda$, Equation (4.15) implies

$$
X^*(t) = \mathcal{H}E_t \left[ \frac{\xi(T)}{\xi(t)} 1_{\xi(T) < \gamma_1} \right] + (1 - K)E_t \left[ \frac{\xi(T)}{\xi(t)} 1_{\gamma_1 \leq \xi(T) < \gamma_2} \right] + E_t \left[ i(\lambda \xi(T)) \frac{\xi(T)}{\xi(t)} 1_{\gamma_1 \leq \xi(T) < \gamma_2} \right],
$$

i.e. Equation (4.9) with

$$
d_1(t) = (\ln(\gamma_1/\xi(t)) - |\vartheta|^2(T - t)/2)/\left(|\vartheta|\sqrt{T - t}\right), \\
d_2(t) = d_1(t) + |\vartheta|\sqrt{T - t}/(1 - a), \\
d_3(t) = (\ln(\gamma_2/\xi(t)) - |\vartheta|^2(T - t)/2)/\left(|\vartheta|\sqrt{T - t}\right), \\
d_4(t) = d_3(t) + |\vartheta|\sqrt{T - t}/(1 - a).
$$

since

$$
E_t \left[ \frac{\xi(T)}{\xi(t)} 1_{\xi(T) < \gamma_1} \right] = N(d_1(t)), \quad E_t \left[ \frac{\xi(T)}{\xi(t)} 1_{\xi(T) < \gamma_2} \right] = N(d_3(t)), \\
E_t \left[ \frac{\xi(T)}{\xi(t)} i \frac{\lambda \xi(T)}{1 + m} 1_{\xi(T) < \gamma_1} \right] = (\mathcal{H} + K - 1) \frac{N'(d_1(t))}{N'(d_2(t))} N(d_2(t)), \\
E_t \left[ \frac{\xi(T)}{\xi(t)} i \frac{\lambda \xi(T)}{1 + m} 1_{\xi(T) < \gamma_2} \right] = (\hat{x} + K - 1) \frac{N'(d_3(t))}{N'(d_4(t))} N(d_4(t)).
$$

The optimal strategy $\delta^*$ is computed as in Appendix 4.A.1 exploiting Equation (4.16).
Chapter 4. On relative performance, remuneration and risk taking of asset managers

4.A.4 Proof of Theorem 4.3.4

In this case \( U(x) := u((x - 1)^+ - (x - \mathcal{H})^+ + mx + K) \) and the concavified utility function

\[
\tilde{U}(x) = \begin{cases} 
U(x) & \text{if } x > \tilde{x}_u, \\
U(0) + U'(\tilde{x}_u)x & \text{if } 0 \leq x \leq \tilde{x}_u, \\
\mathcal{H} & \text{if } x < 0, 
\end{cases} \tag{4.18}
\]

is not differentiable in 0 and \( \mathcal{H} \), therefore we can define a set-valued function \( \tilde{U}' \) as

\[
\tilde{U}'(x) = \begin{cases} 
\{U'(x)\} & \text{if } x > \mathcal{H}, \\
[U'(\mathcal{H}^+), U'(\mathcal{H}^-)] & \text{if } x = \mathcal{H} \\
\{U'(x)\} & \text{if } \tilde{x}_u < x < \mathcal{H}, \\
\{U'(\tilde{x}_u)\} & \text{if } 0 < x \leq \tilde{x}_u, \\
[U'(\tilde{x}_u), +\infty) & \text{if } x = 0. 
\end{cases}
\]

We can also define an inverse for the function \( \tilde{U}' \) given by

\[
I(z) = \frac{1}{1 + m} \left( i \left( \frac{z}{1 + m} \right) + 1 - K \right) 1_{U'(\mathcal{H}^-) < z < U'(\tilde{x}_u)} + \mathcal{H} 1_{U'(\mathcal{H}^+) \leq z \leq U'(\mathcal{H}^-)} \\
+ \frac{1}{m} \left( i \left( \frac{z}{m} \right) + 1 - K - \mathcal{H} \right) 1_{z < U'(\mathcal{H}^+)}.
\]

Let us define \( \gamma_1 = U'(\mathcal{H}^+)/\lambda, \gamma_2 = U'(\mathcal{H}^-)/\lambda, \) and \( \gamma_3 = U'(\tilde{x}_u)/\lambda, \) then

\[
X^*(t) = \frac{1}{1 + m} E_t \left[ \frac{\xi(T)}{\xi(t)} i \left( \frac{\lambda \xi(T)}{1 + m} \right) 1_{\gamma_2 < \xi(T) < \gamma_3} \right] + \frac{1 - K}{1 + m} E_t \left[ \frac{\xi(T)}{\xi(t)} 1_{\gamma_2 < \xi(T) < \gamma_3} \right] \\
+ \mathcal{H} E_t \left[ \frac{\xi(T)}{\xi(t)} 1_{\gamma_1 < \xi(T) < \gamma_2} \right] + \frac{1}{m} E_t \left[ \frac{\xi(T)}{\xi(t)} i \left( \frac{\lambda \xi(T)}{m} \right) 1_{\xi(T) < \gamma_1} \right] \\
+ \frac{1 - K - \mathcal{H}}{m} E_t \left[ \frac{\xi(T)}{\xi(t)} 1_{\xi(T) < \gamma_1} \right],
\]

i.e. Equation (4.10) with

\[
d_1(t) = \left( \ln(\gamma_1/\xi(t)) - |\vartheta|^2(T - t)/2 \right) / \left( |\vartheta| \sqrt{T - t} \right), \\
d_2(t) = d_1(t) + |\vartheta| \sqrt{T - t}/(1 - a), \\
d_3(t) = \left( \ln(\gamma_2/\xi(t)) - |\vartheta|^2(T - t)/2 \right) / \left( |\vartheta| \sqrt{T - t} \right), \\
d_4(t) = d_3(t) + |\vartheta| \sqrt{T - t}/(1 - a), \\
d_5(t) = \left( \ln(\gamma_3/\xi(t)) - |\vartheta|^2(T - t)/2 \right) / \left( |\vartheta| \sqrt{T - t} \right), \\
d_6(t) = d_5(t) + |\vartheta| \sqrt{T - t}/(1 - a).
\]

The optimal strategy \( \delta^* \) is computed as in Appendix 4.A.1 exploiting Equation (4.16).
Appendix 4.A Proofs

4.A.5 Proof of Theorem 4.3.5

In this case $U(x) := u((x-1)^+ - (x-H)^+ + mx + K)$ and the concavified utility function is given by

$$
\bar{U}(x) = \begin{cases} 
U(x) & \text{if } 0 \leq x < \tilde{x}_d \text{ or } x > \tilde{x}_u, \\
U(x_d) + U'(\tilde{x}_u)(x - x_d) & \text{if } \tilde{x}_d \leq x \leq \tilde{x}_u, \\
-\infty & \text{if } x < 0;
\end{cases}
$$

(4.19)

since $\bar{U}$ is not differentiable in 0 and $H$, we can define a set-valued function $\tilde{U}'$ as

$$
\tilde{U}'(x) = \begin{cases} 
\{U'(x)\} & \text{if } x > H, \tilde{x}_u < x < H, \text{ or } 0 < x < \tilde{x}_d, \\
[U'(H^+), U'(H^-)] & \text{if } x = H, \\
\{U'(\tilde{x}_u)\} & \text{if } \tilde{x}_d \leq x \leq \tilde{x}_u, \\
[U'(\tilde{x}_u), +\infty) & \text{if } x = 0.
\end{cases}
$$

The inverse for the function $\tilde{U}'$ is given by

$$
I(z) = \frac{1}{m} \left( i \left( \frac{z}{m} \right) - K \right) 1_{U'(\tilde{x}_u)<z<U'(0)} + \tilde{x}_u 1_{z=U'(\tilde{x}_u)} + \frac{1}{1+m} \left( i \left( \frac{z}{1+m} \right) + 1 - K \right) 1_{U'(H^-)<z<U'(\tilde{x}_u)} + H 1_{U'(H^+)<z\leq U'(H^-)} + \frac{1}{m} \left( i \left( \frac{z}{m} \right) + 1 - K - H \right) 1_{z<U'(H^+)}.
$$

Let us define $\gamma_1 = U'(H^+)/\lambda$, $\gamma_2 = U'(H^-)/\lambda$, $\gamma_3 = U'(\tilde{x}_u)/\lambda$, and $\gamma_4 = U'(0)/\lambda$, then

$$
X^*(t) = \frac{1}{m} E_t \left[ \frac{\xi(T)}{\xi(t)} i \left( \frac{\lambda \xi(T)}{m} \right) 1_{\gamma_3 < \xi(T) < \gamma_4} \right] - \frac{K}{m} E_t \left[ \frac{\xi(T)}{\xi(t)} 1_{\gamma_3 < \xi(T) < \gamma_4} \right] + \frac{1}{1+m} E_t \left[ \frac{\xi(T)}{\xi(t)} i \left( \frac{\lambda \xi(T)}{1+m} \right) 1_{\gamma_2 < \xi(T) < \gamma_3} \right] + \frac{1}{1+m} E_t \left[ \frac{\xi(T)}{\xi(t)} 1_{\gamma_2 < \xi(T) < \gamma_3} \right] + \frac{\lambda E_t \left[ \frac{\xi(T)}{\xi(t)} 1_{\gamma_1 < \xi(T) < \gamma_2} \right] + \frac{1}{m} E_t \left[ \frac{\xi(T)}{\xi(t)} i \left( \frac{\lambda \xi(T)}{m} \right) 1_{\xi(T) < \gamma_1} \right] + \frac{1}{m} E_t \left[ \frac{\xi(T)}{\xi(t)} 1_{\xi(T) < \gamma_1} \right],
$$

i.e. Equation (4.11) with

$$
d_1(t) = \left( \ln(\gamma_1/\xi(t)) - |\vartheta|^2(T-t)/2 \right) / \left( |\vartheta|\sqrt{T-t} \right),
$$

$$
d_2(t) = d_1(t) + |\vartheta|\sqrt{T-t}/(1-\alpha),
$$

$$
d_3(t) = \left( \ln(\gamma_2/\xi(t)) - |\vartheta|^2(T-t)/2 \right) / \left( |\vartheta|\sqrt{T-t} \right),
$$
Chapter 4. On relative performance, remuneration and risk taking of asset managers

\[ d_4(t) = d_3(t) + |\vartheta|\sqrt{T - t}/(1 - a), \]
\[ d_5(t) = (\ln(\gamma_3/\xi(t)) - |\vartheta|^2(T - t)/2)/\left(|\vartheta|\sqrt{T - t}\right), \]
\[ d_6(t) = d_5(t) + |\vartheta|\sqrt{T - t}/(1 - a), \]
\[ d_7(t) = (\ln(\gamma_4/\xi(t)) - |\vartheta|^2(T - t)/2)/\left(|\vartheta|\sqrt{T - t}\right), \]
\[ d_8(t) = d_7(t) + |\vartheta|\sqrt{T - t}/(1 - a). \]

The optimal strategy \( \delta^* \) is computed as in Appendix 4.A.1 exploiting Equation (4.16).

4.A.6 Proof of Theorem 4.3.6

In this case \( U(x) := u(p(x - H_1)^+ + mx - c(H_2 - x)^+ + K) \) and the concavified utility function is defined as in (4.18) and is not differentiable in 0, therefore we define its first order derivative as

\[ \tilde{U}'(x) = \begin{cases} 
\{U'(x)\} & \text{if } x > \hat{x}_u, 
\{U'(\hat{x}_u)\} & \text{if } 0 < x \leq \hat{x}_u, 
[U'(\hat{x}_u), +\infty) & \text{if } x = 0.
\end{cases} \] (4.20)

Its inverse is given by

\[ I(z) = \frac{1}{m + p} \left(i \left(\frac{z}{m + p}\right) - K + pH_1\right) 1_{z < U'(\hat{x}_u)}. \]

Therefore

\[ X^*(t) = \frac{pH_1 - K}{m + p} E_t \left[\frac{\xi(T)}{\xi(t)} 1_{\xi(T) < \gamma}\right] + \frac{1}{m + p} E_t \left[\frac{\xi(T)}{\xi(t)} t \left(\frac{\lambda \xi(T)}{m + p}\right) 1_{\xi(T) < \gamma}\right], \]

with \( \gamma = U'(\hat{x}_u)/\lambda \), i.e. Equation (4.12) with

\[ d_1(t) = (\ln(\gamma/\xi(t)) - |\vartheta|^2(T - t)/2)/\left(|\vartheta|\sqrt{T - t}\right), \]
\[ d_2(t) = d_1(t) + |\vartheta|\sqrt{T - t}/(1 - a). \]

The optimal strategy \( \delta^* \) is computed as in Appendix 4.A.1 exploiting Equation (4.16).

4.A.7 Proof of Theorem 4.3.7

In this case \( U(x) := u(p(x - H_1)^+ + mx - c(H_2 - x)^+ + K) \) and the concavified utility function is defined as in (4.19) and is not differentiable in 0, therefore we define its first order derivative as

\[ \tilde{U}'(x) = \begin{cases} 
\{U'(x)\} & \text{if } 0 < x < \hat{x}_d \text{ or } x > \hat{x}_u, 
\{U'(\hat{x}_u)\} & \text{if } \hat{x}_d \leq x \leq \hat{x}_u, 
[U'(0), +\infty) & \text{if } x = 0.
\end{cases} \] (4.21)
Appendix 4.A Proofs

In this case the inverse for the function $\tilde{U}'$ is given by

$$I(z) = \frac{1}{m + p} \left( i \left( \frac{z}{m + p} \right) + pH_1 - K \right) 1_{z < U'(\hat{x}_u)} + \hat{x}_u 1_{z = U'(\hat{x}_u)}$$

$$+ \frac{1}{m + c} \left( i \left( \frac{z}{m + c} \right) + cH_2 - K \right) 1_{U'(\hat{x}_u) < z < U'(0)}.$$ 

Therefore we have

$$X^*(t) = \frac{1}{m + p} E_t \left[ \frac{\xi(T)}{\xi(t)} i \left( \frac{\lambda \xi(T)}{m + p} \right) 1_{\xi(T) < \gamma_1} \right] + \frac{pH_1 - K}{m + p} E_t \left[ \frac{\xi(T)}{\xi(t)} 1_{\xi(T) < \gamma_1} \right]$$

$$+ \frac{1}{m + c} E_t \left[ \frac{\xi(T)}{\xi(t)} i \left( \frac{\lambda \xi(T)}{m + c} \right) 1_{\gamma_1 < \xi(T) < \gamma_2} \right] + \frac{cH_2 - K}{m + c} E_t \left[ \frac{\xi(T)}{\xi(t)} 1_{\gamma_1 < \xi(T) < \gamma_2} \right],$$

with $\gamma_1 = U'(\hat{x}_u)/\lambda = U'(\hat{x}_d)/\lambda$, and $\gamma_2 = U'(0)/\lambda$, i.e. Equation (4.13) with

$$d_1(t) = (\ln(\gamma_1/\xi(t)) - |\vartheta|^2(T - t)/2) / \left( |\vartheta|\sqrt{T - t} \right),$$

$$d_2(t) = d_1(t) + |\vartheta|\sqrt{T - t}/(1 - a),$$

$$d_3(t) = (\ln(\gamma_2/\xi(t)) - |\vartheta|^2(T - t)/2) / \left( |\vartheta|\sqrt{T - t} \right),$$

$$d_4(t) = d_3(t) + |\vartheta|\sqrt{T - t}/(1 - a).$$

The optimal strategy $\delta^*$ is computed as in Appendix 4.A.1 exploiting Equation (4.16).


Bibliography


Bibliography


