3rd International Conference on Computational Finance (ICCF2019)

Book of abstracts

Iñigo Arregui
José A. García
Carlos Vázquez
editors
Third International Conference on Computational Finance

A Coruña, 8-12 July 2019
Third International Conference on Computational Finance. Book of abstracts
ARREGUI, Iñigo; GARCÍA, José A.; VÁZQUEZ, Carlos (eds.)
A Coruña, 2019

Universidade da Coruña, Servizo de Publicacións

Number of pages: 227
21 × 29,7 cm
Contents: pp. 5–11
Author index: pp. 225–227

Persistent URL: http://hdl.handle.net/2183/23188
CDU: 336:(519.6+51-3)
IBIC: KFF | UDBM | KJK | KCLF

This book is released under a Creative Commons license
BY-NC-SA (Attribution-NonCommercial-ShareAlike) 4.0 International
Contents

Aknowledgements .......................................................... 13
Foreword .................................................................... 15

Plenary lectures ................................................................ 21

René Aid, Dylan Possamaï, Nizar Touzi
Optimal electricity demand response contracting with responsiveness incentives . . . 23

Stéphane Crépey, Romuald Elie, Wissal Sabbagh, Shiqi Song
When Capital is a Funding Source: The XVA Anticipated BSDEs .............................. 24

Griselda Deelstra
On multivariate European option and barrier option pricing in regime-switching models 25

Bruno Dupire
The Perils of Parameterization .............................................. 26

Peter A. Forsyth
Multi-period Mean CVAR Asset Allocation: Is it Advantageous to be Time Consistent? 27

Julien Guyon
The Joint S&P 500/VIX Smile Calibration Puzzle Solved ................................. 28

Monique Jeanblanc, Libo Li
Characteristics of random times .............................................. 29

Ruediger Kiesel, Nikolaus Graf von Luckner, Anke Kramer
Analysis of intraday power markets using Hawkes processes ............................ 30
Andrea Pallavicini
Valuation adjustments in practice ........................................... 31

Olivier Pironneau
Applications of Deep Learning to Partial Differential Equations and Finance .... 32

Christoph Reisinger, Yufei Zhang
Deep neural network approximations to high-dimensional control and games in finance 33

Nizar Touzi
Continuous time Principal Agent and optimal planning .......................... 34

Antony Ware
Polynomial maps of polynomial processes for energy prices .................. 35

Minisymposium 1: Analytical and numerical methods for option pricing ........ 37

Manuel L. Esquível, Nadezhda P. Krasii
Some financial market models obtained by Euler discretization of continuous models 39

João Guerra, Manuel Guerra, Zachary Polaski
Asset allocation using option-implied distributions in an exponentially tempered stable Lévy model ............................................ 43

Sara D. Lopes, Carlos Vázquez
LIBOR market model including credit risk under real world measure ............ 44

Ana. M. Monteiro, Antonio A. F. Santos
Large scale nonparametric estimation of risk-neutral densities through jointly use of constraints based on call and put option prices ........................... 45

Carlos Oliveira
Investment with decreasing cost due to technological innovation improvements ... 50

Michele Pignotti, María Suárez-Taboada, Carlos Vázquez
Mining extraction projects: mathematical analysis and numerical methods for new PDE models ................................................... 51

Minisymposium 2: Approximation methods from numerics, probability and statistical learning in computational finance .......................... 53

Kathrin Glau, Daniel Kressner, Francesco Statti
Low-rank tensor approximation for Chebyshev interpolation in parametric option pricing .............................................................. 55

Michèle Vanmaele
Pricing American options using radial basis functions .............................. 56

Wolfgang Stockinger
Simulation of McKean SDEs, multi-level Monte Carlo schemes and applications in computational finance ........................................ 57

Elisabetta Carlini, Athena Picarelli, Francisco Silva
A truncated semi-Lagrangian scheme for second order degenerate Hamilton-Jacobi-Bellman equations in bounded domains ........................... 58

Damir Filipovic, Amir Khalilzadeh
Statistical learning for higher moments of asset returns .......................... 59
Stefan Gerhold, Martin Forde, Benjamin Smith  
Small-time and large-time smile behaviour for the rough Heston model  

Yufei Zhang  
A penalty scheme and policy iteration for stochastic hybrid control problems with nonlinear expectations  

Lotfi Boudabsa, Damir Filipović  
Learning with kernels in finance  

Linus Wunderlich, Kathrin Glau  
Efficient numerical techniques for parametric problems in option pricing  

Christian P. Fries  
Stochastic Algorithmic Differentiation of discontinuous functions (and other stochastic operators) and application in mathematical finance  

Minisymposium 3: Consumption and investment under mortality risk  

Petar Jevtic, Minsuk Kwak, Traian Pirvu  
Longevity bond pricing in equilibrium  

Mogens Steffensen  
Optimal design: from insurance policy to economic policy  

Fabio Gómez, Jaime Londoño  
Optimal consumption, investment, and life insurance purchase: a state-dependent utilities approach  

Jaime A. Londoño  
State-dependent utility with mortality risk  

Minisymposium 4: Local volatility models and inverse problems  

Jorge P. Zubelli, Vinicius Albani  
A splitting strategy for the calibration of jump-diffusion models  

Bruno Dupire  
A new view at local volatility and local correlation  

Stéphane Crépey  
Deep calibration of local volatility  

Julien Guyon  
Inversion of convex ordering: local volatility does not maximize the price of VIX futures  

Minisymposium 5: Machine learning methods in computational finance  

Sofie Reyners  
Machine learning for quantitative finance: fast derivative pricing  

Lukas Gonon  
Reservoir computing with stochastic inputs: volatility forecasting  

Ganna Marchenko  
Maximum entropy approach for nonstationary time series analysis with application to volatility modelling
Philippe von Wurstemberger
Overcoming the course of dimensionality with DNNs: Theoretical approximation results for PDEs

Minisymposium 6: Modelling and valuation techniques for energy markets
María del Carmen Calvo-Garrido, Matthias Ehrhardt, Carlos Vázquez
Pricing swing options in electricity markets with two stochastic factors: PIDE modeling and numerical solution

Michael Coulon, Jonas Ströjby
Wind park valuation and risk management in German intraday power markets

Klaus Spanderen
Efficient Gas Swing and Storage valuation with Neural network and PDE/MC

Derek Bunn, Ekaterina Abramova
Forecasting intra-day price spread density matrices for electricity storage operations

Minisymposium 7: Numerical Methods for PDEs in finance
Lynn Boen, Karel in ’t Hout
Operator splitting schemes for the two-asset Merton jump-diffusion model

Christina Christara, Yuwei Chen
A penalty-like method for CVA pricing by a PDE model

Olivier Bokanowski, Kristian Debrabant
BDF finite difference schemes for diffusion equations with obstacle

Bertram Düring, Alexander Pitkin
High-order compact finite difference scheme for option pricing in stochastic volatility jump models

Minisymposium 8: Qualitative and quantitative aspects of nonlinear PDEs based models arising in mathematical finance
Iñigo Arregui, Beatriz Salvador, Daniel Ševčovič, Carlos Vázquez
PDE models for American options with total value adjustment and two stochastic factors

Daniel Ševčovič, José Cruz
On solutions of a partial integro-differential Black-Scholes equation in Bessel potential spaces

Soňa Kilianová, Mária Trnovská, Daniel Ševčovič
Dynamic portfolio optimization via a Hamilton-Jacobi-Bellman equation

José Cruz, Daniel Ševčovič
Option pricing in illiquid markets with jumps

Igor Melicherčík, Aleš Černý, Gábor Szücs
Pension saving strategies based on Samuelson’s lifecycle theory

Vera Egorova, Rafael Company, Lucas Jódar
An exponential time differencing method for American option pricing problems
Minisymposium 9: Young researchers minysimposium on computational finance

Emmanuel Gobet, José G. López-Salas, Carlos Vázquez
Quasi-Regression Monte-Carlo scheme for semi-linear PDEs and BSDEs 121

Outhmane Mounjid, Mathieu Rosenbaum, Pamela Saliba
A ranking methodology for market making activity 122

Côme Huré
Some neural network-based schemes to solve high-dimensional nonlinear PDEs 123

Shuoqing Deng, Xin Li, Huyen Pham, Xiang Yu
Optimal relative consumption by competing with past spending maximum 124

Ankush Agarwal, Emmanuel Gobet
Finite variance unbiased estimation of stochastic differential equations 125

Lokman Abbas-Turki
Conditional Monte Carlo learning for diffusions 126

Arnaud Lionnet
Numerical approximation of BSDEs with polynomial growth driver 127

Lorenc Kapllani, Long Teng, Matthias Ehrhardt
Multistep schemes for solving Backward Stochastic Differential Equations on GPU 128

Contributed talks

Jennifer Alonso-García, Michael Sherris, Samuel Thirurajah, Jonathan Ziveyi
Incorporating taxation in the valuation of variable annuity contracts: the case of the guaranteed minimum accumulation benefit 131

Pol Álvarez, Argimiro Arratia
HPC framework for multiple evaluation of supervised learning models and a famous expert for stock screening using fundamental factors 135

Argimiro Arratia, Albert Dorador
Models of stock prices with overnight gap and flash crash to test stop-loss rules efficiency 139

C. Bayer, C. Ben Hammouda, R. Tempone
Hierarchical adaptive sparse grids for option pricing under the rough Bergomi model 143

Nikolay Bondarenko, Geoffrey Kingston, Pavel Shevchenko
Optimal life-contingent insurance under bid-ask spreads 144

Anastasia Borovykh
Understanding generalisation in noisy time series forecasting 146

Dan Cheng, Pasquale Cirillo
A reinforced URN process modeling of recovery rates and recovery times 147

Pasquale Cirillo, Dan Cheng
A urn-based nonparametric modeling of the dependence between PD and LGD with an application to mortgages 149

Javier de Frutos, Víctor Gatón
A spectral method for an optimal investment problem with transaction costs under potential utility 154

Andrea Fontanari, Iddo Eliazar, Pasquale Cirillo, Cornelis W. Oosterlee
Portfolio risk and the quantum majorization of correlation matrices 155

Raquel Gaspar
Empirics on CPPI design risk 157
Lech A. Grzelak, Anton van der Stoep, Kees Oosterlee
Collocating Local Volatility: A competitive alternative to stochastic local volatility models .................................................... 161

Archil Gulisashvili, Raúl Merino, Marc Laguna, Josep Vives
Higher order approximation of call option prices under stochastic volatility models ............................................. 165

Nikolay Gudkov, Jonathan Ziveyi
Pricing and hedging of guaranteed minimum benefits using power series approximation techniques ............................................. 169

Jonathan Hüser, Uwe Naumann
Regression likelihood ratio method Monte Carlo sensitivities ................................................................. 173

Emma M. Iglesias, Garry D. A. Phillips
Pseudo-maximum likelihood estimation and testing in the constant elasticity of variance continuous time model ................................................................. 178

Margaret Insley, Peter A. Forsyth
Climate games: Who’s on first? What’s on second? ................................................................. 179

Kentaro Kikuchi
A term structure interest rate model with the exit time from the Quantitative Easing Policy ................................................................. 180

Fabien Le Floc’h, Cornelis W. Oosterlee
Arbitrage-free implied volatility interpolation by spline collocation ................................................................. 184

Geoffrey Lee, Zili Zhu, Bowie Owens
Introducing two mixing fractions to a lognormal LSV model ................................................................. 190

Sunju Lee, Younhee Lee
Real option pricing on finite time horizon ................................................................. 194

Álvaro Leitao, Luis Ortiz-Gracia, Justin L. Kirkby
Continuous Time Markov Chain approximation of the Heston model ................................................................. 195

Shuaiqiang Liu, Cornelis W. Oosterlee
Calibration neural networks for financial models ................................................................. 200

Zororo S. Makumbe, Youssef El-Khatib, Josep Vives
Some results on the Hybrid Heston-LSV Model With Jumps: Pricing, Hedging and Sensitivities. A Malliavin approach ................................................................. 201

Zsolt Nika
Investment strategies with long memory: log-optimal solution and approximations by stochastic gradient ................................................................. 205

María R. Nogueiras, Ignacio Ruiz, Mariano Zeron
Numerical approximation of dynamic initial margin: Chebychev on the market space analysed ................................................................. 206

Rebeca Peláez, Ricardo Cao, Juan M. Vilar
Probability of default estimation using a nonparametric approach ................................................................. 208

Vladimir Petrov, Anton Golub, Richard Olsen
Instantaneous volatility seasonality of high-frequency markets in directional-change intrinsic time ................................................................. 210

Jiang Pu
Accelerated Share Repurchase: option hedging, optimal execution and neural networks ................................................................. 211

Paulo M. Silva
Robo-advising: unfolding the risks ................................................................. 214
Jacob Snoeijer, Karel J. in ’t Hout
Efficient numerical valuation of high-dimensional basket options via partial differential equations .................................................. 216

Serguei Vassilovski, Alex Kreinin
Multi-curve interest rate modeling with Hull-White and Black-Karasinski ............. 217

Samyukta Venkataramanan, Stephen O’Sullivan
Parameter estimation of affine term structure models ........................................ 221

Guan Yan, Stefan Trück
Interconnectedness of spot electricity prices. A dynamic network analysis ........... 223

Wenjun Zhang, Jiling Cao, Gaurav Kapoor, Xinfeng Ruan
Path-dependent leveraged exchange-traded fund option pricing .......................... 224

Authors index ......................................................................................................... 225
Acknowledgements

The Organizing Committee wants to thank the plenary lecturers, minisimposia organizers, speakers and members of the Scientific Committee. Also our acknowledgements to all the attendants and participants of the present edition of ICCF. All of them highly contributed to the success of ICCF2019 in A Coruña.

We also want to thank the support of the following institutions:

- Abanca
- Afundación
- Centro de Investigación en Tecnologías de la Información y Comunicaciones (CITIC)
- European Consortium for Mathematics in Industry, ECMI
- Fundación Pedro Barrié de la Maza
- Grupo de investigación M2NICA de la Universidad de A Coruña
- Horizon 2020 EU Program, through the MSCA-ITN-EID "ABC-EU-XVA"
- Instituto Tecnológico de Matemática Industrial, ITMATI
- Universidade da Coruña
- Xunta de Galicia, through the Program "Ayudas de Consolidación para Grupos de Referencia Competitiva"

The Organizing Committee
A Coruña, 8 July 2019
This volume contains the abstracts of the Third International Conference on Computational Finance (ICCF 2019), held in A Coruña, Spain, between 8 and 12 July 2019, and coorganized by the Universidade de A Coruña, http://iccf2019.udc.es/.


Initiated in 2015, this is the third of a biennial series of conferences in Computational Finance launched by the consortium of the European project FP 7 Marie Curie ITN STRIKE-Novel Methods in Computational Finance 2013-2016. Former editions took place in December 2015 at University of Greenwich (United Kingdom) and in September 2017 at Universidade de Lisboa (Portugal). As a key activity of the ECMI Special Interest Group on Computational Finance and Energy Markets, the ICCF 2019 was held in A Coruña.

These series of conferences represent a forum for financial industry scientists, experts and researchers, to present recent and breakthrough developments in modelling and numerical methods in financial problems, and also to promote the collaboration between academia and financial industry.

This edition focused on recent advances in research and practice in financial mathematics and computing, for both academic researchers and practitioners. It consisted of plenary lectures, given by international leading academics and industry researchers working on subjects of particular relevance for the theory and practice of quantitative finance, and contributed talks organized in thematic sessions.

The purpose of ICCF2019 is:

- To foster discussions and collaborations among applied mathematicians, probabilists, statisticians, computational specialists and researchers and practitioners in financial industry, deepening cooperation and promoting a proficuous interplay with the cross-fertilization of ideas.

- To provide an opportunity for advanced students to develop competences on financial mathematics and computing by contacting with specialists both from academia and industry.
The Special Interest Group proposed several sessions in Mini-Symposia at the ICCF 2019. Topics that are included are:

- Stochastic Analysis and Control Theory in Finance
- Lévy Processes in Finance
- Interest Rate Modelling
- Credit Risk Modelling
- Portfolio Management
- Models for Risk Management
- Commodities Markets
- Energy Market Models
- Big Data Analytics
- Computational Methods and High Performance Computing in Finance
- Market Microstructure.

Finally, it should be pointed that more than 100 professors, researchers, students and industrial experts attended and/or presented their achievements, thus contributing to the success ICCF2019.

The Organizing Committee
A Coruña, 8 July 2019
Scientific committee

Michael Coulon, University of Sussex, UK
Matthias Ehrhardt, Bergische Universität Wuppertal, Germany
Maria Rosario Grossinho, ISEG Lisboa, Portugal
Choi-Hong Lai, University of Greenwich, UK
Jaime A. Londoño, Universidad Nacional de Colombia
María R. Nogueiras, HSBC London, UK
Cornelis W. Oosterlee, CWI and TU Delft, The Netherlands
Andrea Pascucci, Università di Bologna, Italy
Daniel Ševčovič, Comenius University, Slovakia
Carlos Vázquez, Universidade da Coruña, Spain (Chair)
Jorge Zubelli, IMPA, Brazil
Organizing committee

Ana Almécija, Universidade da Coruña
Iñigo Arregui, Universidade da Coruña (Chair)
Mª Carmen Calvo, Universidade da Coruña
Ana Mª Ferreiro, Universidade da Coruña
José A. García, Universidade da Coruña
Álvaro Leitao, Universidade da Coruña
José G. López, Universidade da Coruña
Beatriz Salvador, Universidade da Coruña
María Suárez, Universidade da Coruña
Carlos Vázquez, Universidade da Coruña
Plenary lectures
Optimal electricity demand response contracting
with responsiveness incentives

RENÉ AÏD
UNIVERSITÉ PARIS–DAUPHINE, PSL UNIVERSITY
EMAIL: rene.aid@dauphine.psl.eu
URL: https://sites.google.com/view/reneaid/home

DYLAN POSSAMAÏ
COLUMBIA UNIVERSITY
EMAIL: dp2917@columbia.edu

NIZAR TOUZI
ÉCOLE POLYTECHNIQUE
EMAIL: touzi@polytechnique.edu
URL: http://www.cmap.polytechnique.fr/touzi/

Abstract. Despite the success of demand response programs in retail electricity markets in reducing average consumption, the literature shows failure to reduce the variance of consumers’ responses. This paper aims at designing demand response contracts which allow to act on both the average consumption and its variance. The interaction between the producer and the consumer is modeled as a Principal-Agent problem, thus accounting for the moral hazard underlying demand response programs. The producer, facing the limited flexibility of production, pays an appropriate incentive compensation in order to encourage the consumer to reduce his average consumption and to enhance his responsiveness. We provide closed-form solution for the optimal contract in the case of linear energy valuation. Without responsiveness incentive, this solution decomposes into a fixed premium for enrolment and a proportional price for the energy consumed, in agreement with previously observed demand response contracts. The responsiveness incentive induces a new component in the contract with payment rate on the consumption quadratic variation. The calibration of our model to publicly available data of a large scale demand response experiment predicts a significant increase of responsiveness under our optimal contract, a significant increase of the producer satisfaction, and a significant decrease of the consumption volatility.
Abstract. XVAs refer to various financial derivative pricing adjustments accounting for counterparty risk (CVA) and its funding (FVA) and capital (KVA) implications for a bank. We show that the XVA equations are well posed, including in the realistic case where capital (including capital at risk) is deemed fungible as a source of funding for variation margin. The intertwining of capital at risk and FVA, added to the fact that the KVA is part of capital at risk, then lead to a system of anticipated McKean BSDEs (ABSDEs) for the FVA and the KVA, with coefficients entailing a conditional risk measure of the one-year-ahead increment of the martingale part of the FVA. But these ABSDEs can be recast as a forward/backward SDE system, which suggests a Picard iteration, shown convergent, for decoupling the forward simulation of the trading P&L of the bank (martingale part of the CVA plus the FVA, in the converging limit of the Picard iteration) from the backward XVA and economic capital pricing task.
Abstract. First, we will concentrate upon the pricing of some multivariate European options when the underlying asset prices are driven by Markov-modulated Lévy processes (MMLPs) with dependence between the asset’s components, and where synchronous jumps are allowed. In particular, the asset prices may jump whenever there is a change of phase of the underlying Markov process. When considering exchange options and quanto options, pricing formulae are based upon the characteristic exponents by using the well known FFT methodology. For the evaluation of spread and basket options in this general regime-switching framework, we provide lower and upper bounds to the exact option prices based upon ideas from Caldana and Fusai (2013) and Caldana et al. (2016). These bounds are obtained via univariate Fourier inversion under the assumption that the joint characteristic functions of the MMLPs are known.

Next, we consider the risk-neutral pricing of vanilla, digital and down-and-out call options when the (single) underlying asset price evolves like the exponential of a Markov-modulated Brownian motion (MMBM) with two-sided phase-type jumps. The price of such options is intimately related to the first passage properties of the MMBM. To analyse these first passages, we randomize the time horizon using Erlang distributions with suitable parameters and apply matrix-analytic methods. This provides us with closed form approximations of the option prices, with a very high precision, as shown by several numerical illustrations.

Joint work with: Sinem Kozpinar, Guy Latouche, Matthieu Simon.
Abstract. Markets trade thousands of underlyings, each one with tens or even hundreds of options, quoted throughout the day. Needless to say, the quotes are not generated manually, they are automated and derived from a functional form with a few parameters. If we know this parameterization, we know in advance that the prices tomorrow of many traded securities will belong to a low dimensional (number of parameters) manifold in a high dimensional (number of securities). If the vector of today prices does not belong to the convex hull of the manifold it creates arbitrage. We examine market practice (Black-Scholes, stochastic volatility models, interest rate interpolation by piecewise constant instantaneous forward rates converging implied volatility for extreme strikes in FX...) and show that many violate the no arbitrage condition.
Multi-period Mean CVAR Asset Allocation: Is it Advantageous to be Time Consistent?

PETER A. FORSYTH*
CHERITON SCHOOL OF COMPUTER SCIENCE, UNIVERSITY OF WATERLOO, WATERLOO ON, CANADA, N2L 3G1
EMAIL: paforsyt@uwaterloo.ca
URL: https://cs.uwaterloo.ca/paforsyt/

Abstract. We formulate the multi-period, time consistent mean-CVAR (Conditional Value at Risk) asset allocation problem in a form amenable to numerical computation. Our numerical algorithm can impose realistic constraints such as: no shorting, no-leverage, and discrete rebalancing. We focus on long term (i.e. 30 year) strategies, which would be typical of an investor in a Defined Contribution (DC) pension plan. A comparison with pre-commitment mean-CVAR strategies shows that adding the time consistent constraint compares unfavourably with the pure pre-commitment strategy. Since the pre-commitment strategy computed at time zero is identical to a time consistent strategy based on an alternative objective function, the pre-commitment mean-CVAR strategy is implementable in this case. Hence it would seem that there is little to be gained from enforcing time consistency.

Keywords: multi-period mean-CVAR, time consistent, pre-commitment, asset allocation

*This work was supported by the Natural Sciences and Engineering Research Council of Canada
The Joint S&P 500/VIX Smile Calibration Puzzle Solved

JULIEN GUYON
BLOOMBERG L.P., QUANTITATIVE RESEARCH, 731 LEXINGTON AVE, NEW YORK, NY 10022, USA
EMAIL: jguyon2@bloomberg.net
URL: http://cermics.enpc.fr/~guyon/

Abstract. Since VIX options started trading in 2006, many researchers and practitioners have tried to build a model that jointly and exactly calibrates to the prices of S&P 500 (SPX) options, VIX futures and VIX options. So far the best attempts, which used continuous-time jump-diffusion models on the SPX, could only produce an approximate fit. In this talk we solve this puzzle using a discrete-time model. Given a VIX future maturity $T_1$, we build a joint probability measure on the SPX at $T_1$, the VIX at $T_1$, and the SPX at $T_2 = T_1 + 30$ days which is perfectly calibrated to the SPX smiles at $T_1$ and $T_2$, and the VIX future and VIX smile at $T_1$. Our model satisfies the martingality constraint on the SPX as well as the requirement that the VIX at $T_1$ is the implied volatility of the 30-day log-contract on the SPX. In particular, this proves that the SPX and VIX markets are jointly arbitrage-free. The discrete-time model is cast as a dispersion-constrained martingale transport problem and solved using the Sinkhorn algorithm, in the spirit of De March and Henry-Labordère (2019). We explain how to handle the fact that the VIX future and SPX option monthly maturities do not perfectly coincide, and how to extend the two-maturity model to include all available monthly maturities.
Characteristics of random times

Monique Jeanblanc
LaMME, 23 Boulevard de France, 91037 Evry Cedex, France
Email: monique.jeanblanc@univ-evry.fr
URL: http://www.maths.univ-evry.fr/pages_perso/jeanblanc/

Libo Li
UNSW, Sydney
Email: libo.li@unsw.edu.au

Abstract. The purpose of this talk is to present the processes associated to a random time (we call them characteristics) useful to perform computations: pricing, portfolio optimization. One of them is the intensity, another one is the Azéma supermartingale, a last one is the conditional density.

In a second part, we show how we can construct random times with given characteristics.
Analysis of intraday power markets using Hawkes processes

Ruediger Kiesel
Universität Duisburg-Essen, Germany
Email: Ruediger.Kiesel@uni-due.de

Nikolaus Graf von Luckner
Universität Duisburg-Essen, Germany
Email: nikolaus.graf-von-luckner@uni-due.de

Anke Kramer
Universität Duisburg-Essen, Germany
Email: Anke.Kramer@uni-due.de

Abstract. Due to the increasing share of renewable energies the German intraday electricity market (IDM) has gained importance in recent years. The trading activity on the IDM, in particular, the arrival of market orders, can be modeled by point processes. Intraday market data provide evidence that orders tend to arrive in clusters. Therefore, self-exciting processes are often used to account for this feature. We focus on whether Hawkes processes with time-varying baseline intensities and exponential kernels are suited to model these counting processes. We deploy standard methods in the context of counting processes to assess goodness-of-fit. We apply our model to study the problem of optimal market making on the IDM.
Valuation Adjustments in Practice

ANDREA PALLAVICINI
DEPARTMENT OF MATHEMATICS, IMPERIAL COLLEGE, LONDON.
EMAIL: a.pallavicini@imperial.ac.uk

Abstract. The financial crisis started in 2007 has shown that any pricing framework must include from the very beginning the possibility of default of any market player. As a consequence, derivative valuation and risk analysis have moved from exotic derivatives managed on simple single-asset classes to simple derivatives embedding credit risk and new, or previously neglected, types of complex and interconnected non-linear effects.

The detailed analysis of how a trade is really implemented between two counterparties in the market is a requirement to track all the possible costs and risks. Derivative valuation is adjusted to include counterparty credit risk and contagion effects along with funding costs due to collateral posting, treasury policies, and regulatory constraints. A second level of complexity is produced by moving from a single trade to the whole bank portfolio. Aggregation-dependent valuation processes, and theirs operational challenges, arising from non-linearities are discussed both from a mathematical and practical point of view.

By following the developments of recent literature, an arbitrage-free valuation framework is developed for bilateral counterparty risk adjustments, inclusive of wrong-way and contagion risks, collateralization and funding costs. The resulting valuation equations take the form of semi-linear PDEs, or backward SDEs, depending only on real market rates and processes, and no longer on unobservable risk-free rates. Collateralization is described starting from the day-by-day operations followed by the counterparties to implement the credit support annex. Segregation and re-hypothecation policies are discussed both for bilateral-traded and centrally-cleared deals along with funding implications. Margining and funding costs are included into the derivation of pricing formulae, along with their relationship with hedging strategies and liquidity policies. Fund transfer pricing procedures driving the business strategies of the bank are presented and their impact on funding costs is described.
Applications of Deep Learning to Partial Differential Equations and Finance

Olivier Pironneau
Sorbonne University, Paris, France
Email: olivier.pironneau@upmc.fr

Abstract. Neural networks have been game changers in many optimization and classification problems. Google, Facebook and Nvidia have written highly optimized and user-friendly open source toolboxes with which it is easy to assess to potentials of deep learning on control, inverse and classification problems for distributed systems. On the other hand the mathematical results which certify that the problems are solved are few.

In this talk we will try to survey the mathematical results available and present applications on 3 classes of problems:

- Classification of linear PDEs.
- Identification of parameters for fluid structure problem and for the Heston model in finance.
- Computation of solutions of linear parabolic PDE in high dimensions.
Deep neural network approximations to high-dimensional control and games in finance

CHRISTOPH REISINGER
UNIVERSITY OF OXFORD, UNITED KINGDOM
EMAIL: christoph.reisinger@maths.ox.ac.uk

YUFEI ZHANG
UNIVERSITY OF OXFORD, UNITED KINGDOM
EMAIL: yufei.zhang@maths.ox.ac.uk

Abstract. In this talk, we discuss the feasibility of algorithms based on deep artificial neural networks (DNN) for the solution of high-dimensional stochastic control problems and zero-sum games arising in financial engineering. In the first part, we show that in certain cases, including those with open loop controls, the value functions can be represented by a DNN whose complexity grows at most polynomially in both the dimension of the state equation and the reciprocal of the required accuracy, demonstrating that DNNs can break the curse of dimensionality. In the second part, we exploit policy iteration to reduce the nonlinear problem into a sequence of linear problems, which are then further approximated via a multilayer feedforward neural network ansatz. We establish that in suitable settings the numerical solutions converge globally in the $H^2$-norm. Moreover, we construct the optimal feedback controls based on the superlinear convergence of the numerical solutions. Preliminary numerical experiments are presented to illustrate the theoretical results and to demonstrate the effectiveness of the method.
Abstract. Using representation results from second order backward SDEs, we provide a general solution approach to continuous time Principal Agent problems. The same technique allows also to formulate the optimal planning problem in mean field games as an optimal transport problem along controlled dynamics.
Polynomial maps of polynomial processes for energy prices

ANTONY WARE
DEPT. OF MATHEMATICS AND STATISTICS, UNIVERSITY OF CALGARY, CANADA
EMAIL: aware@ucalgary.ca
URL: people.ucalgary.ca/~aware

Abstract. In the context of energy price modelling, prices are formed from exponential maps of underlying factor processes, and the mathematical convenience this offers means that this is no surprise. In this talk we will show various ways in which models based on polynomial maps of polynomial processes (PMPP models) can function in a similar way.

Polynomial processes have the property that expectations of polynomial functions of the future state of the process, conditional on the current state, are themselves polynomial functions of the current state. It is this property that means that PMPP models also provide a level of mathematical convenience (for forming futures prices). But they also provide an additional level of flexibility, which means that they are capable of capturing the extreme dynamics that are commonly seen in energy market prices even with relatively tame dynamics in the underlying factor process.

We will end by discussing numerical methods for the valuation of energy contracts in the PMPP setting.
Minisymposium 1:
Analytical and numerical methods for option pricing

Organized by
Maria do Rosário Grossinho
Some financial market models obtained by Euler discretization of continuous models

MANUEL L. ESQUIVEL
DEPARTMENT OF MATHEMATICS AND CMA, FCT NOVA, UNL, 2829–516 CAPARICA, PORTUGAL
EMAIL: mle@fct.unl.pt
URL: http://ferrari.dmat.fct.unl.pt/personal/mle/

NADEZHDAA P. KRASHI
DON STATE TECHNICAL UNIVERSITY, GAGARIN SQUARE 1, ROSTOV-ON-DON 344000, RUSSIAN FEDERATION
EMAIL: krasnad@yandex.ru
URL: https://donstu.ru/structure/cadre/krasiy-nadezhda-pavlovna

Abstract. We present and study several discrete time financial models with one risky asset and a risk free asset that may thought to result as a discretization of a suitable continuous time model. We compare the pricing results, obtained with these models, with results obtained from the related continuous time models. Our approach relies on some known important results describing a particular class of discrete time models – the conditionally Gaussian models – a class that, nevertheless, contains many interesting instances. We aim at a better understanding of the implications of the discretization procedures which are inevitable, both at the parameter estimation and derivative price computation moments, by reason of the computational implementations.

1. Introduction. Given a financial market model with dynamics defined by stochastic differential equations (SDE) we have that often, the estimation procedures rely in the observation of the processes at discrete times and so, it would be natural to consider discrete time models associated to the initially given continuous time models. As it happens, the Euler-Maruyama discretization scheme (see [3, p. 62]) allows us to consider models for which it is possible to retrieve a computable martingale measure. The case of a SDE model with deterministic coefficients, dealed in Section 3 below, is easier. A more general case, with the coefficients being allowed to be random, is treated in Section 4.

2. Discretization of a discounted price process. Let \((\Omega, \mathcal{F}, \mathbb{P})\) be a complete probability space and \((B_t)_{t \geq 0}\) a standard Brownian motion and \(\mathcal{F} = (\mathcal{F}_n)_{n \geq 0}\) the Brownian filtration. For the financial market model we may consider a SDE of the type:

\[
dS_t = \mu_t S_t dt + \sigma_t S_t dB_t, \quad t \in [0, T], \quad S_0 = S_0 \in \mathbb{R}_+^*,
\]

as a model for stock prices and admitting a strong solution. We suppose that the risk free rate process \((\rho_t)_{t \in [0,T]}\) of the market satisfies \(d\rho_t = \rho_t dt, \quad t \in [0,T], \quad \rho_0 = 1\) and that \((\mu_t)_{t \in [0,T]}, (\sigma_t)_{t \in [0,T]}\) and \((\rho_t)_{t \in [0,T]}\) are predictable with respect to \(\mathcal{F}\). Then (see [2, pp. 160]), the discounted price process \((\tilde{S}_t)_{t \in [0,T]}\) given by,

\[
d\tilde{S}_t = (\mu_t - \rho_t) \tilde{S}_t dt + \sigma_t \tilde{S}_t dB_t, \quad t \in [0, T], \quad S_0 = S_0 \in \mathbb{R}_+^*, \quad \tilde{S}_0 = \exp \left( \int_0^t \left( \mu_s - \rho_s - \frac{\sigma_s^2}{2} \right) ds + \int_0^t \sigma_s dB_s \right).
\]
And so, we have that the process \((X_t)_{t \in [0,T]}\) defined by:

\[ X_t = \log \left( \tilde{S}_t \right), \]

verifies the SDE given by,

\[ dX_t = \left( \mu_t - \rho_t - \frac{\sigma_t^2}{2} \right) dt + \sigma_t dB_t. \tag{2} \]

We now consider the Euler-Maruyama discretization of a stochastic differential equation having in mind applying it to an equation such as (2). Let \( T < +\infty \) be the temporal time horizon, and for a given integer \( N \geq 1 \) let:

\[ \Delta t := \Delta^N t = \frac{T}{N}. \]

Suppose that \((X_t)_{t \in [0,T]}\) is a continuous time stochastic process adapted to Brownian filtration \((\mathcal{F}_t)_{t \in [0,T]}\). Consider the discretized stochastic process given by:

\[ X_{k\Delta t}, \ k = 0, 1, \ldots, N, \]

that has a natural interpretation as the process of observations of the continuous time process made at the epochs \(0, \Delta t, 2\Delta t, \ldots, T\). Let us consider the discrete time stochastic process defined by:

\[ \forall k \geq 0, \ X_{(k+1)\Delta t} = X_{k\Delta t} + \mu \left( (k+1)\Delta t \right) \Delta t + \sigma \left( (k+1)\Delta t \right) \Delta t \cdot Z_{k+1}, \tag{3} \]

where we have that \((Z_k)_{k \geq 1}\) is a sequence of standardized normal random variables.

3. The non random coefficients case. In the simpler case — for formula (3) — in which \((\mu_t)_{t \in [0,T]}, (\sigma_t)_{t \in [0,T]}\) are non random, we have that \(\mathbb{F} = (\mathcal{F}_n)_{n \geq 0}\) may be the natural filtration generated by \((Z_k)_{k \geq 0}\) or, equivalently by \((X_{k\Delta t})_{k \geq 0}\) — as the coefficients \(\mu, \sigma\) are non random — and we have the following result.

**Proposition 3.1.** The sequence of random variables defined by,

\[ Y_k = \exp \left( -\sum_{n=0}^{k} \frac{\mu (n\Delta t)}{\sigma (n\Delta t)} Z_n - \sum_{n=0}^{k} \frac{\mu (n\Delta t)^2}{2\sigma (n\Delta t)^2} \right), \]

is an \(\mathbb{F}\)-martingale and the process \((X_{k\Delta t} Y_k)_{n \geq 0}\) is an \(\mathbb{F}\)-martingale.

As a consequence, by considering the probability space \((\Omega, \mathcal{F}, \mathbb{Q})\) with \(\mathbb{Q}\) defined by:

\[ \forall F \in \mathcal{F} \quad \mathbb{Q}[F] = \int_F Y_N d\mathbb{P}, \]

also represented by \(d\mathbb{Q} = Y_N d\mathbb{P}\) — with \(Y_N\) the Radon-Nicodym density of \(\mathbb{Q}\) relatively to \(\mathbb{P}\) — we have that the discretized process is a \(\mathbb{Q}\) martingale with respect to \(\mathbb{F} = (\mathcal{F}_n)_{n \geq 0}\) that is,

**Proposition 3.2.** In the space \((\Omega, \mathcal{F}, \mathbb{Q})\) the process \((X_{k\Delta t})_{k=0,1,\ldots,N}\) is an \(\mathbb{F}\) martingale.

**Remark 3.3.** We observe that we may now easily compute the arbitrage free prices of derivatives in a financial market model with discounted prices given by formula (2) in the case where the coefficients are deterministic. For that it is obviously required to estimate the coefficients (see for instance [7]).
4. The random coefficients case. For the case where the coefficients of both the price SDE model and the risk free rate are random we may also consider the discrete market model obtained by the Euler-Maruyama discretization of the continuous market model by using what we may call the Girsanov theorem in discrete time in the conditionally Gaussian case. In [8, pp. 433–446] and in [1, pp. 123–125] we may find a detailed explanation that can be summarized in the next result.

**Theorem 4.1.** Let \((\Omega, \mathcal{F}, \mathbb{P})\) be a complete probability space and \(\mathcal{F} = (\mathcal{F}_t)_{t \geq 0}\) a filtration over this space such as \(\mathcal{F}_0 = \{\emptyset, \Omega\}\). Let the process \(X = (X_n)_{n \geq 1}\) verify the following decomposition:

\[
\forall n \geq 0, \quad X_{n+1} = X_n + \mu_{n+1} + \sigma_{n+1}Z_{n+1},
\]

with \((\mu_n)_{n \geq 1}\) and \((\sigma_n)_{n \geq 1}\) previsible processes, \((Z_n)_{n \geq 1}\) an \(\mathcal{F}\)-adapted process of independent random variables identically distributed such that the law of \(Z_n\) conditioned to \(\mathcal{F}_{n-1}\) is a standardized normal random variable. Let the Girsanov process be given by:

\[
Y_n := \exp \left( -\sum_{k=1}^{n} \frac{\mu_k}{\sigma_k} - \frac{1}{2} \sum_{k=1}^{n} \left( \frac{\mu_k}{\sigma_k} \right)^2 \right), \quad n \geq 1
\]

such that the following Novikov condition is verified:

\[
\forall n \geq 1, \quad \mathbb{E}\left[ \frac{1}{2} \sum_{k=1}^{n} \left( \frac{\mu_k}{\sigma_k} \right)^2 \right] < +\infty.
\]

We then have that:

1. the Girsanov process \(Y = (Y_n)_{n \geq 1}\) is a \(\mathcal{F}\) uniformly integrable martingale which converges, almost surely to a random variable \(Y_\infty\) verifying:

\[
\mathbb{P}[Y_\infty \geq 0] = 1, \quad \mathbb{E}[Y_\infty] = 1;
\]

2. considering the probability measure \(\mathbb{Q}\) defined over \((\Omega, \mathcal{F})\) by:

\[
\forall F \in \mathcal{F}, \quad \mathbb{Q}[F] = \int_F Y_\infty d\mathbb{P},
\]

then, in the probability space \((\Omega, \mathcal{F}, \mathbb{Q})\), the probability law of \((X_{n+1} - X_n)_{n \geq 1}\) conditioned by \(\mathcal{F}\), coincides, term by term, with the probability law of \((\sigma_nZ_n)_{n \geq 1}\);

3. in the probability space \((\Omega, \mathcal{F}, \mathbb{Q})\), the process \((X_{n+1} - X_n)_{n \geq 1}\) is a local martingale.

We observe that the hypothesis made that \(Z_n\) conditioned to \(\mathcal{F}_{n-1}\) is a standardized Gaussian implies that the probability law of \(X_{n+1} - X_n\) conditioned to \(\mathcal{F}_{n-1}\) is given for \(F \in \mathcal{F}\), by:

\[
\mathbb{P}\left[ (X_{n+1} - X_n) \in F \mid \mathcal{F}_{n-1} \right] := \mathbb{E} \left[ \mathbb{1}_{(X_{n+1} - X_n) \in F} \mid \mathcal{F}_{n-1} \right] = \frac{1}{\sqrt{2\pi\sigma_n^2}} \int_F e^{-\frac{(y-\mu_n)^2}{2\sigma_n^2}} dy.
\]

Following [8, pp. 62, 103] this formula may be interpreted by saying that the probability law of \(X_{n+1} - X_n\) conditioned to \(\mathcal{F}_{n-1}\) is a mixture of Gaussian random variables with parameters \(\mu_n\), and \(\sigma_n^2\) that in themselves are random variables

**Remark 4.2.** The second result in the theorem 4.1. is that in the probability space \((\Omega, \mathcal{F}, \mathbb{Q})\) the law of \(X_{n+1} - X_n\) conditioned to \(\mathcal{F}_{n-1}\) is given for \(F \in \mathcal{F}\), by:

\[
\mathbb{Q}\left[ (X_{n+1} - X_n) \in F \mid \mathcal{F}_{n-1} \right] := \mathbb{E}_\mathbb{Q} \left[ \mathbb{1}_{(X_{n+1} - X_n) \in F} \mid \mathcal{F}_{n-1} \right] = \frac{1}{\sqrt{2\pi\sigma_n^2}} \int_F e^{-\frac{y^2}{2\sigma_n^2}} dy,
\]

that is, with the new probability \(\mathbb{Q}\), the coefficient \(\mu_n\) is suppressed and we now have a local martingale.
5. Discrete conditionally Gaussian models. The methodology proposed in this work may now be described as follows. We have discrete time observations of a discounted price process and we want to model it as a discrete conditionally Gaussian model to which we may apply theorem 4.1. in order to obtain a martingale measure and to price derivatives. We may consider following [8, pp. 89, 104–108], a discrete model in which the discounted price process \((S_n)_{n \geq 1}\), defined in a filtered probability space \((\Omega, \mathcal{F}, \mathbb{P}, \mathbb{F})\) and adapted to \(\mathbb{F} = (\mathcal{F}_n)_{n \geq 0}\) satisfying,

\[
S_n = S_0 e^{h_1 + h_2 + \ldots h_n} \text{ with } \mathbb{E}[|h_n|] < +\infty ,
\]

(6)

with \(h_n = \mu_n + \sigma_n \epsilon_n\) where \((\epsilon_n)_{n \geq 1}\) is a sequence of independent standardized Gaussian random variables adapted to \(\mathbb{F}\) and,

\[
\text{Law}(h_n | \mathcal{F}_{n-1}) = N(\mu_n, \sigma_n^2) ,
\]

(7)

with the processes \((\mu_n)_{n \geq 1}\) and \((\sigma_n)_{n \geq 1}\) predictable with respect to \(\mathbb{F}\) and formula (7) having an interpretation similar to the one given for formula (5). This class of models for \((h_n)_{n \geq 1}\) is quite rich as it encompasses, among others, the ARMA\((p,q)\), the ARCH\((p)\) and the GARCH\((p,q)\) models. For instance, for the ARMA\((p,q)\) model we have \(\mathcal{F}_n = \sigma(\epsilon_1, \ldots, \epsilon_n)\), the prescribed initial conditions \(h_{-p}, \ldots, h_{-1}, h_0\) and \(\epsilon_{-q}, \ldots, \epsilon_{-1}, \epsilon_0\) and,

\[
\mu_n = a_0 + a_1 h_{n-1} + \ldots + a_p h_{n-p} + b_1 \epsilon_{n-1} + \ldots + b_p \epsilon_{n-q} ,
\]

with \(\sigma_n \equiv \sigma\) constant. Estimating the adequate time series model we may get under suitable hypothesis a martingale measure and the possibility of computing derivative prices. There are nevertheless important questions to be dealt with, namely, under what conditions, the model given by (6) and (7),

\begin{itemize}
  \item is a Euler-Maruyama discretized model of a continuous time model? That is, does the discretized model converges—see [6]—to some continuous time market model with good properties?
  \item is a arbitrage free and complete model—see [4] and [5]—for all discretization steps considered?
\end{itemize}

We will develop some instances of models studied under the methodology here presented.

References


Asset allocation using option-implied distributions in an exponentially tempered stable Lévy model

JOÃO GUERRA
CEMAPRE, REM and ISEG, Universidade de Lisboa, Portugal
Email: jguerra@iseg.ulisboa.pt

MANUEL GUERRA
CEMAPRE, REM and ISEG, Universidade de Lisboa, Portugal
Email: mguerra@iseg.ulisboa.pt

ZACHARY POLASKI
ISEG, Universidade de Lisboa, Portugal
Email: zack.polaski1@gmail.com

Abstract. We explore the fundamental optimal asset allocation problem using option-implied distributions when the underlying risky asset price is modeled by an exponentially tempered stable Lévy process. The risk-neutral densities are extracted from option prices and are transformed to the risk-adjusted (real-world) densities. The optimal portfolios are constructed, rebalanced periodically and their performance is analyzed. We found that the portfolios formed using option-implied distributions under the Lévy market model, which are flexible enough to capture the higher moments of the implied distribution, are far more robust to left-tail market risks and offer statistically significant improvements to risk-adjusted performance when investor risk aversion is low, however this diminishes as risk aversion increases.

Keywords: Asset Allocation, Lévy processes, Option-Implied Distributions, Portfolio Optimization

JEL Classification: C10, C51, G11, G13, G17
LIBOR Market Model including credit risk under the real world measure

SARA DUTRA LOPES
UNIVERSIDADE DE LISBOA, ISEG, CEMAPRE, PORTUGAL
Email: sblopes@iseg.ulisboa.pt

CARLOS VÁZQUEZ
UNIVERSIDADE DA CORUÑA, CAMPUS DE ELVIÑA, s/n, 15071 A CORUÑA, SPAIN
Email: carlosv@udc.es

Abstract. In this article, we present a methodology to simulate the evolution of interest rates for multiple credit ratings under the real world probability measure, as needed for risk management calculations, e.g., to compute Value at Risk indicators.

Given the current market conditions where negative interest rates are present, in our methodology the presence of negative interest rates is allowed while ensuring the positivity of credit spreads, taking into account that investors demand a risk premium for bonds with higher credit risk.

More precisely, we implement the multidimensional Shifted Lognormal LIBOR market model with credit risk ratings and specify the the market price of risk vector process needed to perform simulations of the real world forward rates in future.

Keywords: Real World Model, Scenario Simulation, Interest Rate, Shifted Lognormal Market Model, Market Price of Risk, Credit Risk, Credit Ratings

References


Large scale nonparametric estimation of risk-neutral densities through jointly use of constraints based on call and put option prices

A. M. Monteiro
Faculdade Economia, CeBER, GEMF, Universidade Coimbra, Av. Dias da Silva, 165, Coimbra, Portugal
Email: ammm@fe.uc.pt

Antonio A. F. Santos
Faculdade Economia, CeBER, GEMF, Universidade Coimbra, Av. Dias da Silva, 165, Coimbra, Portugal
Email: aasantos@fe.uc.pt

Abstract. Option pricing theory determines the structure of call and put option pricing functions. Usually, in a nonparametric risk neutral density estimation based on kernel functions, constraints are imposed locally, not allowing for restrictions related to the fact that density integrates one, and to global imposed convexity and monotonicity. Also, we propose a new global risk-neutral density estimation approach, based on kernel functions imposing those constraints and permitting the simultaneously use of calls and puts without applying the put-call parity formula. Our approach is tested through a simulation experiment using Heston’s model.

1. Introduction. The availability of huge amounts of option data leads to different ways of dealing with existing problems. By this way new approaches and methods could appear giving rise to new strategies and opportunities. A few years ago a fundamental change occur in capital markets allowing access to massive intra-day high-frequency data. In fact, among other factors, electronic trading and high-frequency trades are responsible for this change in available data. This huge amount of data, usually called Big Data, plays a fundamental role in several domains namely, in risk management. However, this process raises some concerns since it augments the susceptibility to errors.

We propose a novel approach for risk-neutral density (RND) estimation based on a nonparametric regression setting [2, 1, 7]. First, we consider a strategy to solve a generalized least squares problem by estimating simultaneously all variables of the problem. This contrast with existing kernel approaches that follow a local nonparametric estimation. Global estimation allows to redefine a set of constraints related to the definition of a density function and to no-arbitrage constraints. Regarding these last constraints, existing approaches using kernel functions do not provide a guarantee that, density estimates integrate one over its support. Second, following [6] we use call and put option prices directly in the optimization problem. Previous approaches have used one-side version of option prices, since they essentially used call prices through put-call parity formula. Using call and put option prices simultaneously improves the robustness of RND estimation. Reformulating our local polynomial approach [6], we propose a nonparametric kernel method using global optimization.

2. No-arbitrage constraints and risk-neutral densities. [4] and [3] derive a way of estimating RNDs from financial options prices, by considering second derivatives of option pricing functions. The European call option price, for contracts on the underlying asset $S$, can be given by

$$C(S_t, X, \sigma_t, \tau, r, \delta) = e^{-r\tau} \int_X^{\infty} (S_T - X) f(S_T | S_t, \sigma_t, \delta, \tau) dS_T,$$

where $t$ represents the current date, $X$ the strike price, $r$ the risk-free interest rate, $S_t$ the current underlying asset price, $S_T$ the asset price at maturity, $\tau = T - t$ the time-to-maturity, $\delta$ the dividend
yield, and $f(\cdot)$ represents the conditional risk-neutral density for the underlying asset at expiration $T$. When considering the price of a European put option, we have

$$P(S_t, X, \sigma_t, \tau, \delta, \tau) = e^{-r\tau} \int_0^X (X - S_T) f(S_T|S_t, \sigma_t, \tau, \delta, \tau) dS_T.$$  

Central to the results presented in this paper is the fact that, risk-neutral density functions and respective cumulative distribution functions are inter-related through first and second derivatives of call and put pricing functions. They can be expressed as,

$$f(S_T|S_t, \sigma_t, \tau, \delta, \tau) = e^{r\tau} \frac{\partial^2 C(S_t, X, \sigma_t, \tau, r, \delta)}{\partial X^2}|_{X=S_T} = e^{r\tau} \frac{\partial^2 P(S_t, X, \sigma_t, \tau, r, \delta)}{\partial X^2}|_{X=S_T} \quad (1)$$

$$F(S_T|S_t, \sigma_t, \tau, \delta, \tau) = 1 + e^{r\tau} \frac{\partial C(S_t, X, \sigma_t, \tau, r, \delta)}{\partial X}|_{X=S_T} = e^{r\tau} \frac{\partial^2 P(S_t, X, \sigma_t, \tau, r, \delta)}{\partial X^2}|_{X=S_T}. \quad (2)$$

These relations between first and second derivatives of call and put pricing functions, risk-neutral densities and respective cumulative distribution functions will be explored within the nonparametric estimation process. Through a no-arbitrage framework, bound constraints for the call pricing function and respective derivatives assume the form

$$-e^{-r\tau} \leq \frac{\partial C}{\partial X}(X) \leq 0; \quad \frac{\partial^2 C}{\partial X^2}(X) \geq 0; \quad C(X) \geq 0, \quad \forall X \in [0, \infty[. \quad (3)$$

By using put-call parity, the same kind of constraints can be associated with the put pricing function

$$0 \leq \frac{\partial P}{\partial X}(X) \leq e^{-r\tau}; \quad \frac{\partial^2 P}{\partial X^2}(X) \geq 0; \quad P(X) \geq 0, \quad \forall X \in [0, \infty[. \quad (4)$$

These constraints, and the equality ones above, will be used within a quadratic optimization problem associated with an extended criterion function. The constraints just considered are the only ones possible in the implementation of a sequential local estimation approach. If a global framework is considered, a further set of constraints can be added, namely the ones that impose that the call (put) pricing function is monotone decreasing (increasing) on strikes, and respective first derivatives are monotone increasing, implying the convexity of the respective pricing functions. The aim is to estimate a risk-neutral density, as non-negativity constraints were already considered, we only have to add the constraint related to integration to one. This will be the subject mainly addressed in this paper, and made us develop what we designated as a global nonparametric estimation.

3. Global vs local nonparametric estimation. When applied to an economical problem, theoretical results can give different degrees of guidance on sound characteristics of a function relating dependent and independent variables. The cases where the bounds too stringent are rare, meaning that, for example, the theory says that on average when one variable grows the other also grows, and vice-versa. [2] presented a nonparametric kernel regression problem to estimate RNDs. For the kernel regression, it is assumed that the process can be described as $y_i = m(x_i) + \sigma(x_i)e_i$, where $m(x)$ is the unknown function to be estimated at a given point $x$. If the function needs to be approximated for different points, in the sense of approximating the function as a whole, sequential local nonparametric estimations can be performed. [2] use the Nadaraya-Watson (NW) estimator, where the implied volatility function is defined through a kernel estimator. Black-Scholes formula is then used to calculate a call pricing function and its second derivative, the estimated risk neutral density. Other authors such as [1, 9, 8] improve this approach by introducing new features to the respective optimization problem.

When a more rich structure is given by the theory, its inclusion in the estimation process will certainly be useful in terms of amelioration of estimators’ statistical properties. This is clearly the case of estimation of risk-neutral density estimation through option prices. The RND corresponds to a function that is defined through the second derivative of option pricing functions, in this context represented by $m(x)$, where $x$ represents strike prices.
Theoretical results establish a set of important constraints on the possible forms of option pricing functions, and also of their first and second derivatives. For call pricing functions, as functions of strikes, it must be assured that they are monotone decreasing and convex. Also near the boundaries, they must comply with no-arbitrage constraints. On the other hand, their first derivatives resembles a cumulative distribution function subtracted by one, and the second derivatives must be proportional to a density function. This is a lot of structure that should be incorporated within an estimation method to obtain interpretable estimates. The same kind of analysis can be performed for puts, considering monotone increasing and convex functions. These important characteristics cannot be taken into account when a sequential point by point estimation method, without global constraints, is implemented to approximate call and put pricing functions, and respective derivatives. Without such global constraints, there is no guarantee that estimated functions are monotone or convex. On the other hand, as the main object of interest is the risk-neutral density, and due to the necessary characteristics of a density function, positiveness and integration to one, the global estimation approach guarantees that estimates comply with such characteristics.

In the context of RND estimation through nonparametric regression methods, using NW estimator is equivalent to a zero order local polynomial approximation (local constant), the associated optimization problem is an unconstrained one. We begin with a local cubic polynomial approximation, considering jointly calls and puts, and the restrictions from (1) to (4). In this paper, we went further by considering global nonparametric estimation method that approximates the objects of interest, for a set of points \( \{x_j\}_{j=1}^m \) representing a partition for the range of RND to estimate.

With this general approach we can include constraints implicit on (1) through (4), constraints associated with monotonicity and convexity of option pricing functions, and also, the one making explicit that estimated RNDs must integrate one. Theoretical results establish a set of important constraints on the possible forms of option pricing functions, and also of their first and second derivatives. These elements are encompassed on \( \beta_j^{(k)}(x_j) \), which are estimated jointly. This framework makes use of all the theoretical structure behind the definition of risk-neutral prices and respective densities associated with options. All the drawbacks associated with a partial view of the problem, making use of cleaning data procedures, smoothing intermediate
Large scale nonparametric estimation of risk-neutral densities

Figure 1: Heston option prices with put and call mean estimates (left-side); RND estimation; sequential local estimation (red), median (solid line), 90% confidence intervals (dashed lines); global estimation (blue), median (solid line), 90% confidence intervals (dashed lines)

steps, and even ad-hoc tuning procedures can be removed by this global view of the problem. It has to be noticed that this approach can result in a large scale optimization problem. Considering that the functions of interest are approximated through 100 points, the decision vector has length of 800, and 1601 constraints need to be considered. As the quadratic programming problem is strictly convex it has a feasible solution, having the advantage of being able to integrate all relevant factors that need to be considered.

4. Empirical demonstration. Our aim in this research is to demonstrate that RNDs estimation accuracy is improved using a global nonparametric estimation. This approach is compared with a sequential local nonparametric estimation. The former allows a new set of constraints to be included in the optimization problem associated. The global approach allows to consider all no-arbitrage constraints in the estimation process. Using this general framework, pricing functions will comply with theoretical-based results, and the resulting RND estimation gives a proper density function, strictly positive that integrates one. Compared with a sequential local approach, more accurate estimates are obtained, which results in narrower confidence intervals.

To demonstrate the feasibility and robustness of the considered global nonparametric estimation of the RND, we considered a simulated experiment where the [5] model is used, in this case the true RND is known. By simulating \( n \) samples with Heston’s prices, added with errors, the estimation process is applied successively and the distribution of the estimators can be approximated, which allows the computation of confidence intervals. Using [5] model allows us to consider a more general diffusion process which is subjected to stochastic volatility, but still compatible with a closed form for option prices. By considering a negative value for the leverage effect, negative asymmetric RNDs are obtained. For the underlying process it is assumed 

\[
\frac{dS_t}{S_t} = \mu dt + \sqrt{v_t} dW_t, \quad dv_t = \kappa(\theta - v_t) dt + \sigma \sqrt{v_t} dZ_t,
\]

where \( W_t \) and \( Z_t \) are two standard Brownian motion processes with \( E(dW_t dZ_t) = \rho dt \), \( \kappa \) represents the mean-reverting volatility parameter, \( \theta \) the long-run volatility, and \( \sigma \) the volatility of volatility. The parameters adopted are \( \kappa = 5 \), \( \theta = 0.03 \), \( \sigma = 0.3 \), and \( \rho = -0.7 \). In the simulation a zero dividend yield, a risk-free rate \( r = 0.02 \), and a time to maturity of 3-months (\( \tau = 0.25 \)), are adopted. It is assumed that at \( t \) the price of the underlying asset is \( S_t = 50 \), and the range of strikes is given by the interval \([35, 62]\). For prices obtained using the model, random noise was added to mimic observed market prices.
In Figure 1 is depicted the results of a simulation using the model described above. Given the model and respective parameters, theoretical prices for calls and puts are calculated, and using a fine grid for strike prices, first and second derivatives are calculated numerically with high precision. This is assumed to be the true RND. For the range of strikes considered, steps of 0.5 were adopted, and for each, 79 observations were simulated for calls and puts, mimicking intraday data with a frequency of 5-min within a daily transaction period. In the figure it is compared the true RND and respective estimates given through the evolution of three quantiles, 5%, 50%, and 95%. Using the extreme quantiles, a confidence interval of 90% is established, and we can appreciate the degree of accuracy obtained. In the tails, intervals are extremely narrow, less in the center due to the curvature, and its influence on the bias and standard error of nonparametric estimators. However, due to the structure imposed, mean and first derivatives compatible with no-arbitrage constraints, and proper density functions are a by-product of the method, without the need of further intermediate steps within the estimation process.

5. Concluding remarks. The approach developed in this paper leads to a proper smooth risk-neutral density due to the imposed constraints considered in a global estimation framework. With this approach no further ad-hoc tuning procedures need to be implemented in order to obtain interpretable estimates. Through the definition of a large scale optimization problem, including all relevant constraints, it is guaranteed that estimates mimic monotone and convex option pricing functions. Second derivatives estimates are compatible with the definition of a RND, positiveness and integration to one. Further, simulation results show that this approach represents an improvement in terms of accuracy when compared with a sequential local estimation framework.

References


Investment with decreasing cost due to technological innovation improvements

CARLOS OLIVEIRA  
ISEG, UNIVERSITY OF LISBON, PORTUGAL  
EMAIL: oliveira.cmds@gmail.com

Abstract. In this paper we address, in the context of real options, an investment problem with two sources of uncertainty: the price (reacted in the revenue of the rm) and the level of technology. The level of technology impacts in the investment cost, that decreases when there is a technology innovation. The price follows a geometric Brownian motion, whereas the technology innovations are driven by a Poisson process. As a consequence, the investment region may be attained in a continuous way (due to an increase of the price) or in a discontinuous way (due to a sudden decrease of the investment cost). For this optimal stopping problem no analytical solution is known, and therefore we propose a quasi-analytical method to find an approximated solution that preserves the qualitative features of the exact solution. This method is based on a truncation procedure and we prove that the truncated solution converges to the solution of the original problem.

We provide results for the comparative statics for the investment thresholds. These results show interesting behaviors, particularly, the investment may be postponed or anticipated with the intensity of the technology innovations and with their impact on the investment cost.
Mining extraction projects: mathematical analysis and numerical methods for new PDE models

MICHELE PIGNOTTI
DEPARTMENT OF MATHEMATICS, UNIVERSITY OF BOLOGNA, ITALY
Email: michele.pignotti2@unibo.it

MARÍA SUÁREZ-TABOADA
DEPARTMENT OF MATHEMATICS, UNIVERSITY OF A CORUÑA, SPAIN
Email: mariasuarez@udc.es

CARLOS VÁZQUEZ
DEPARTMENT OF MATHEMATICS, UNIVERSITY OF A CORUÑA, SPAIN
Email: carlosv@udc.es

Abstract. In this work, a new mathematical model related to a mining extraction project under uncertainty is proposed. We mainly follow the modelling approaches developed in [1] and [3], where the commodity price and the resource size are considered as the relevant stochastic factors. The main innovative modelling feature comes from considering the extraction rate to be proportional to the commodity price. In this way, an ultraparabolic hypoelliptic differential operator governs the associated PDE of the mathematical model. The mathematical analysis allows to obtain the existence and uniqueness of a classical solution. Uniqueness follows from a suitable Feynman-Kac representation formula and existence is obtained after a suitable change of variables, the determination of sub and supersolutions and passing to the limit from problems in bounded domains to the unbounded one. For the numerical solution, after justifying the required boundary conditions on the computational bounded domain, the proposed numerical techniques consist of a Crank-Nicolson characteristics method for the time discretization to cope with the convection dominating setting and Lagrange finite elements for the discretization in the commodity and resource variables. Finally, some numerical examples are discussed. Further details are available in [2].

References
Minisymposium 2: Approximation methods from numerics, probability and statistical learning in computational finance

Organized by
Kathrin Glau
Low-rank tensor approximation for Chebyshev interpolation in parametric option pricing

KATHRIN GLAU*
EPF LAUSANNE AND QUEEN MARY UNIVERSITY OF LONDON
EMAIL: k.glau1@qmul.ac.uk
URL: http://www.maths.qmul.ac.uk/ kglau/

DANIEL KRESSNER
EPF LAUSANNE
EMAIL: daniel.kressner@epfl.ch

FRANCESCO STATTI
EPF LAUSANNE

Abstract. We propose a new method for the efficient computation of high-dimensional parametric conditional expectations such as option prices. The method extends the recently developed approach of [1] to high dimensionality in the parameter space. To overcome the curse of dimensionality, we additionally exploit low-rank structures. The core of our method is to express the tensorized interpolation in tensor train (TT) format and to develop an efficient way, based on tensor completion, to approximate the interpolation coefficients. The method is based on a novel adaptive sampling strategy for tensor completion. Efficiency is, moreover, gained from an offline/online decomposition.

As examples we consider American option pricing in the Heston model and European basket option pricing in the multi-dimensional Black-Scholes model. We treat parameter spaces of dimensions up to 25. The numerical results confirm the low-rank structure of these problems and the effectiveness of our method compared to advanced techniques.

Figure 1: Representation of financial data. Left: tensor of values at interpolation nodes, middle: reduction by sampling, right: data representation in TT-format.

References


*This project has received funding from the European Union’s Horizon 2020 research and innovation program under grant 665667.)
Pricing American options using radial basis functions

MICHÈLE VANMAELE
DEPARTMENT OF APPLIED MATHEMATICS, COMPUTER SCIENCE AND STATISTICS, GHENT UNIVERSITY, BELGIUM
EMAIL: Michele.Vanmaele@UGent.be

Abstract. We study the price of American options as solutions to partial differential equations. We will discuss several methods depending on the type of option and the type of model. For example under Heston’s stochastic volatility model we propose a radial basis function with partition of unity method applied to a linear complementary formulation of the free boundary partial differential equation problem. This results in an algebraic problem with sparse matrices which have a moderate condition number. Next, a Crank-Nicolson time discretisation is combined with the operator splitting method to get a fully discrete problem. To better control the computational cost and the accuracy, adaptivity is used in the spatial discretisation.

Numerical experiments are performed with European and American options, and results are compared with numerical results available in the literature to illustrate the accuracy and efficiency of the proposed algorithms.
Abstract. In the first part of this talk, we will present schemes for the simulation of McKean SDEs, e.g., the particle method which is used to approximate the true distribution of the current state of some process. Considering time-stepping schemes, classical approaches, such as the Euler-Maruyama method and the tamed Euler-scheme find applications for these SDEs. Here, we also introduce an adaptive scheme and demonstrate its performance numerically for several examples.

Moreover, we discuss a range of existing Multi-level Monte Carlo schemes for McKean SDEs and consider their applications in computational finance. In particular, calibration conditions for local stochastic volatility models (LSVMs) lead to a complex class of McKean SDEs, which require the use of kernel density estimation methods. Quasi-Monte Carlo methods applied to this problem can significantly improve the variance convergence rate of estimators for certain conditional expectations appearing in the calibration problem. This will be demonstrated numerically for multiple examples of LSVMs.
A truncated semi-Lagrangian scheme for second order degenerate Hamilton-Jacobi-Bellman equations in bounded domains

Elisabetta Carlini ∗
Piazzale Aldo Moro 5, Roma, Italy
Email: carlini@mat.uniroma1.it

Athena Picarelli†
Via Cantarane 24, Verona, Italy
Email: athena.picarelli@univr.it

Francisco Silva‡
123, Avenue Albert Thomas, Limoges, France
Email: francisco.silva@unilim.fr

Abstract. We study the numerical approximation of parabolic, possibly degenerate, Hamilton-Jacobi-Bellman (HJB) equations in bounded domains. It is well known by the seminal paper [1] that convergence of the numerical approximation to the exact solution of the equation (considered here in the viscosity sense) is achieved under the assumptions of monotonicity, consistency and stability of the scheme. While standard finite difference schemes are in general non monotone, the so-called semi-Lagrangian (SL) schemes (see [2, 3] and the references therein) are monotone by construction. These schemes make use of a wide stencil and, when the equation is set in a bounded domain, this typically causes an overstepping of the boundary. A truncated version of SL schemes has been presented in [4]. We discuss here a suitable modification of this scheme adapted to the treatment of degenerate boundary problems.

References

∗University of Rome La Sapienza
†University of Verona
‡University of Limoges
Statistical Learning For Higher Moments of Asset Returns

DAMIR FILIPOVIC  
SWISS FINANCE INSTITUTE AND EPF LAUSANNE  
EMAIL: damir.filipovic@epfl.ch

AMIR KHALILZADEH  
SWISS FINANCE INSTITUTE AND EPF LAUSANNE  
EMAIL: amir.khalilzadeh@epfl.ch

Abstract. Statistical learning methods has received increasing attention in financial economics recently. In particular, researchers tend to revisit the familiar and fundamental empirical problems such as return prediction and measuring risk premia in asset pricing. We augment this growing literature by investigating the second moment of risky asset returns. In particular, we study the predictability of realized variances and covariances of stock returns for the firms in the US. Knowledge about the direction as well as size of comovements of the asset returns is of great importance to both academia and industry for asset management and portfolio risk management purposes. We do so by designing a statistical learning framework where a large set of stock-level and macro-finance features form a function beyond what is devised by standard linear methods. In particular we deploy decision tree learning method in order to carve the functional form that results in the best out of sample forecast for the realized covariance. Our main finding is that statistical learning methods consistently enrich our understanding of the behaviour of asset return covariances. Ultimately, we are interested in conducting a comparative analysis across various statistical learning methods and find independent key factors that best explain the cross-section and time series of asset returns covariances.
Small-time and large-time smile behaviour for the rough Heston model

Stefan Gerhold
TU Wien
Email: sgerhold@fam.tuwien.ac.at

Martin Forde
King’s College London

Benjamin Smith
King’s College London

Abstract. We characterize the asymptotic small-time and large-time implied volatility smile for the rough Heston model introduced by El Euch, Jaisson and Rosenbaum. We show that the asymptotic short-maturity smile scales in qualitatively the same way as a general rough stochastic volatility model, and is characterized by the Fenchel-Legendre transform of the solution a Volterra integral equation (VIE). The solution of this VIE satisfies a space-time scaling property which simplifies its computation. We corroborate our results numerically with Monte Carlo simulations. We also compute a power series in the log-moneyness variable for the asymptotic implied volatility, which yields tractable expressions for the vol skew and convexity, thus being useful for calibration purposes. We also derive formal asymptotics for the small-time moderate deviations regime and a formal saddlepoint approximation for call options in the large deviations regime. This goes to higher order than previous works for rough models, and in particular captures the effect of the mean reversion term. In the large maturity case, the limiting asymptotic smile turns out to be the same as for the standard Heston model, for which there is a well known closed-form formula in terms of the SVI parametrization.
A penalty scheme and policy iteration for stochastic hybrid control problems with nonlinear expectations

YUEFEI ZHANG  
UNIVERSITY OF OXFORD, UNITED KINGDOM  
EMAIL: yuefei.zhang@maths.ox.ac.uk

Abstract. We propose a penalty method for mixed optimal stopping and control problems where the objective is evaluated by a nonlinear expectation. The solution and free boundary of an associated HJB variational inequality are constructed from a sequence of penalized equations, for which the penalization error is estimated. The penalized equation is then discretized by a class of semi-implicit monotone approximations. We further propose an efficient iterative algorithm with local superlinear convergence for solving the discrete equation. Numerical experiments are presented for an optimal investment problem under ambiguity to demonstrate the effectiveness of the new schemes. Finally, we extend the penalty scheme and its convergence analysis to stochastic hybrid control problems involving continuous and impulse controls.
Learning with kernels in finance

LOTFI BOUDA BSA
ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE, SWITZERLAND
EMAIL: lotfi.boudabsa@epfl.ch

DAMIR FILIPOVIĆ

Abstract. We use Reproducing Kernel Hilbert Spaces to approximate functions in $L^2(E, \mathcal{E}, \rho(dx))$, where $(E, \mathcal{E}, \rho(dx))$ is a probability space. We give both approximation and estimation convergence results (Central Limit Theorems, $L^2$ convergence and concentration inequalities). Furthermore, we show that our approximation/estimation scheme is robust with respect to the function to approximate.

Our scheme has applications in finance. We show how to price and hedge path-dependent options. More generally, our scheme can be seen as an alternative to efficiently compute conditional expectations.
Efficient numerical techniques for parametric problems in option pricing

Linus Wunderlich
Queen Mary University of London
Email: l.wunderlich@qmul.ac.uk
URL: http://www.linuswunderlich.de

Kathrin Glau
École polytechnique fédérale de Lausanne, Queen Mary University of London
Email: k.glau@qmul.ac.uk
URL: http://www.maths.qmul.ac.uk/ kglau/

Abstract. This talk presents some recent developments of efficient numerical techniques, in particular for the fast evaluation of the implied volatility. The implied volatility is one of the most frequently used function in finance and its efficient computation is of high interest in industry. In (Glau, Herold, Madan, and Pötz 2017), a new method to efficiently compute the implied volatility based on multivariate Chebyshev interpolation was introduced. We present adaptations of the method for a high performance implementation of the method.

References

Stochastic Algorithmic Differentiation of discontinuous functions (and other stochastic operators) and application in mathematical finance

CHRISTIAN P. FRIES
EMAIL: email@christian-fries.de

Abstract. In this paper we present a method for the accurate estimation of the derivative (aka. sensitivity) of expectations of functions involving an indicator function by combining a stochastic algorithmic differentiation and a regression.

The method is an improvement of the approach presented in [1, 2].

The finite difference approximation of a partial derivative of a Monte-Carlo integral of a discontinuous function is known to exhibit a high Monte-Carlo error. The issue is obvious, since the Monte-Carlo approximation of a discontinuous function is just a finite sum of discontinuous functions and as such not even differentiable.

The algorithmic differentiation of a discontinuous pay-off is problematic. A natural approach is to replace the discontinuity by continuous functions (aka. pay-off smoothing). This is equivalent to replacing a path-wise automatic differentiation by a (local) finite difference approximation.

We show that this local finite difference approximation can be seen as a linear regression with the simplest regression basis function (a single indicator). Investigating the expression, we observe that we can separate the expectation of the indicator function (the density) and the regression of the size and speed of the discontinuity. With this formulation, we then replace the regression(s) by more accurate estimators.

References


Minisymposium 3: Consumption and investment under mortality risk

Organized by
Jaime Londoño
Longevity bond pricing in equilibrium

PETAR JEVTIC
SCHOOL OF MATHEMATICAL AND STATISTICAL SCIENCES, ARIZONA STATE UNIVERSITY,
TEMPE, AZ 85287-1804, USA

MINSUK KWAK
DEPARTMENT OF MATHEMATICS, HANKUK UNIVERSITY OF FOREIGN STUDIES,
GYEONGGI-DO, 17035, S. KOREA

TRAIAN A. PIRVU
DEPARTMENT OF MATHEMATICS AND STATISTICS, McMASTER UNIVERSITY, ON, L8S 4K1, CANADA

Abstract. Though it is experiencing steady progress in its evolution (see Blake et al. (2018)), the market for mortality linked instruments or so-called life market, is far from reaching its full potential estimated to be of tens of trillions of dollars (see Michaelson and Mulholland (2014)). Currently highly illiquid and compared to equity markets with relatively very low number of transactions, the life market is comprised of series of negotiated, high monetary value, over-the-counter transactions between few agents that have different risk preferences. To accommodate these realities we consider a partial equilibrium model for pricing a longevity bond. We do this under the assumption of stochastic mortality intensity that affects the income of economic agents who trade in risky financial security and longevity bond to maximize their monetary utilities. As a practical contribution in answering an important open question of pricing of longevity bonds in life market, we find the endogenous equilibrium bond price which is numerically computed. In a realistic setting of two agents in a transaction, numerical experiments confirm the expected intuition of price dependence of model parameters.
1. Introduction. We discuss how utility optimization can help to shape improved product design that appropriately unveils and takes into account the demand for life insurance among young individuals and the demand for annuities among old individuals. Classes of objectives lead to classes of optimal life-cycle profiles of consumption, investment, insurance and annuitization decisions that can help pension funds to better target the needs of an individual, both in the savings and the payout phase. To some extent these profiles can also help politicians make an improved design of the pension system as such. We discuss some of the homogeneous issues and dilemmas that turn up in pension policy making in different countries with heterogeneous pension systems. Can utility optimization really help, not only in insurance policy making but also in economic policy making?

2. The dynamics of consumption, investment and insurance. Richard [3] solved the consumption-investment-insurance problem of an individual with uncertain lifetime. The problem and the result were unnoticed for 30 years until actuarial researchers during the 2000s realized its impact. They have then generalized, interpreted, and implemented the ideas during the last decade. Often the consumption, investment and insurance decisions are specified as explicit functions of ages and wealth. With the optimal wealth plugged in, we derive and discuss the dynamics of all three decision processes to learn details about their life-cycle profiles. Many of the stylized facts are well-known although not always rooted in the Richard [3] formulation of the problem. Others are less-known but have appealing connotations, in particular for the actuary.

3. Reasons for Hump-shaped consumption. In positive economics, one aims to explain observed or realized economic behavior with features within a model. Observed hump-shaped consumption patterns are difficult to explain, but the literature does offer a couple of explanations. We present here three: The change of the sign of the market price of diversifiable insurance risk over the lifecycle, as obtained by Richard (1975); lack of exasticity of consumption as alive or dead in combination with patience of consumption, see [1]; and habit formation in combination with impatience for consumption, see [2]. The impact of hump-shaped consumption on the design of pension schemes is obvious. E.g. as decreasing life annuity benefit matches best the demand for consumption. Yet, many countries do not allow for tax benefits of such a product. We discuss this disproportion between individual demand and political supply.

4. Stochastic mortality in a complete market. One generalization of the Richard [3] model is in the direction of systematic (non-diversifiable) mortality risk. Results in that direction are briefly reviewed. Generally, one needs mortality-linked derivatives to complete the market in order to obtained closed-form stylized results. Intuitively, this is easiest to realize because the value of labor income is not unique unless the market has established a market price of systematic mortality risk. An obvious discussion to take in that relation is whether the mortality-linked derivatives market will ever come and why/why not?

5. Stochastic mortality in an incomplete market. In the payout-phase one does not need a mortality-linked derivatives market to complete the market because the labor income is zero. This allows for the calculation of optimal consumption-investment-insurance in a realistic situation where the policy holder has to bear non-diversifiable mortality risk himself. We derive and analyse the solution, in particular the generalization of the dynamics of optimal consumption, investment, and insurance arising from systematic mortality risk. The profile of these dynamics serve as a cornerstone that every pension
scheme designer, provider and law maker must relate to. All deviations from such a profile, being in terms of product design or economic policy (incentives) must be carefully argued for.

References


Optimal consumption, investment, and life insurance purchase: a state-dependent utilities approach

FABIO ANDRÉS GÓMEZ DE LOS RÍOS
Kr 30 No. 45-03, Universidad Nacional de Colombia, Bogotá, Colombia
Email: fagomezde@unal.edu.co

JAIME ALBERTO LONDOÑO LONDOÑO
Kr 27 No. 64-60, Universidad Nacional de Colombia, Manizales, Colombia
Email: jaimelondonol@unal.edu.co

Abstract. We consider the problem of an individual who has to make decisions (under uncertainty) about optimal consumption, investment, and life insurance purchase in a financial market with a finite number of securities; the role of the life insurance is to protect the individual’s family of an eventually early death. We propose facing the problem of optimal election under the alternative approach of state–dependent utilities; so we assume that preferences measure the agent’s satisfaction for future cash flows valued by the market when the individual is making his/her decisions.

1. Introduction. The problem of optimal consumption and investment for a “small investor” whose actions do not influence market prices is at the core of portfolio management. The modern treatment of this problem when asset prices follow Itô processes started with the seminal works of [14, 15]. Merton developed his well-known model solving the problem of optimal consumption and investment under uncertainty but without having into account life insurance acquisition; to obtain explicit solutions, it is necessary that the utility function belongs to the family of HARA (Hyperbolic absolute risk aversion) utility functions. Subsequently, [20] generalized the work of Merton to incorporate life insurance decisions for a consumer whose lifetime is random, independent of the filtration generated by asset prices, and supported on a bounded interval. Richard assumed a finite number of securities whose prices are generated by a geometric Brownian motion; parameters such as labor income, the appreciation rate, and the volatility are deterministic functions. In Richard’s framework, the life insurance is offered instantaneously; the individual buys an insurance that promises to him/her a payment of $p(t)/\lambda(t)$ in case of death at time $t$, the individual pays premiums at the rate $p(\cdot)$ at each time, and the function $\lambda(\cdot)$ is specified in the insurance contract. Here, the individual’s objective is to maximize

$$E \left( \int_0^\tau U_1(s, c(s)) \, ds + U_2(\tau, Z(\tau)) \right),$$

where $\tau \in [0, T]$ is the uncertain lifetime, $c(\cdot)$ is the consumption process, $Z(\cdot)$ is the legacy he/she leaves at death, and $U_1(\cdot, \cdot), U_2(\cdot, \cdot)$ are his/her time-dependent preferences for consumption and bequest, respectively. Richard solved this problem by using dynamic programming techniques; to obtain explicit solutions, he assumed that the utilities belong to the family of CRRA (Constant relative risk aversion) utility functions. Moreover, Richard imposed the condition that the individual cannot purchase the life insurance just at time $T$ if the individual is alive at that time. However, his solution shows that the individual tends to purchase the infinity life insurance just before his maximal possible lifetime $T$. Then, [18] and [23] extended the model proposed by Richard. In this setting, the individual’s lifetime is a random variable which takes values in the interval $(0, \infty)$ and is independent of the process defining the underlying market; now the time $T$ is interpreted as the retirement time. In the model of Pliska and Ye, the individual is endowed with an initial wealth, and he/she also receives a deterministic income continuously until the time of death or retirement (the minimum of them). Here, the individual’s
Optimal consumption, investment, and life insurance purchase

objective is to maximize

\[ E \left( \int_0^{T \wedge \tau} U_1(s, c(s)) \, ds + U_2(Z(\tau), \tau) 1_{\{\tau < T\}} + U_3(X(T)) 1_{\{\tau \geq T\}} \right), \]

where \( X(\cdot) \) represents the wealth process, and \( U_3(\cdot) \) is the utility function for the terminal wealth. Pliska and Ye employed the dynamic programming approach to attack this problem and found explicit solutions for the family of CRRA utility functions. The model of Pliska and Ye relies on the hypothesis that parameters governing the financial market (such as the volatility; and the interest rate) and labor income are deterministic functions.

Following the same line, [7] supposed that the stochastic income is correlated with the risky asset; they modeled the utility of the family instead of separating consumption and bequest and studied numerical solutions. For their part, [5, 4] extended the work of Pliska and Ye by allowing the financial market to have an arbitrary finite number of risky securities whose prices are driven by a multi-dimensional Brownian motion. Shortly after, [17] investigated the problem by considering a complete financial market with one risky security having mean reverting drift and constant volatility, and one riskless security evolving with a constant return per unit of time. Also, the authors modeled labor income as a geometric Brownian motion adapted to the filtration generated by the stock price, and they obtained closed solutions. Then, [21] examined the optimization problem in a complete market; they assumed that the interest rate, the appreciation rate, the volatility, the force of mortality, the premium-insurance ratio, the income, and the discount rate are random processes adapted to the Brownian motion filtration which not necessarily are bounded nor have Markovian structures. In spite of Shen and Wei allowed unbounded random parameters, the interest rate; the force of mortality; the premium-insurance ratio; the discount rate; and labour income must be bounded away from zero, and satisfy some exponential integrability conditions. Later, [3] studied the impact of the introduction of a class of time-inconsistent utilities (heterogeneous discount) in the problem. Recently, [6] generalized the results of [21] to a jump-diffusion model with random parameters. All works cited after [18], with the exception of [3], assume CRRA utility functions to obtain explicit solutions.

Motivated by the structure of pension funds, some authors such as [16] and [10] have considered a variation of the problem proposed by [18]; there, the individual’s objective is to maximize

\[ E \left( \int_0^{T \wedge \tau} U_1(s, c(s)) \, ds + U_2(D(\tau), \tau) 1_{\{\tau < T\}} + U_3(X(T)) 1_{\{\tau \geq T\}} \right), \]

where \( D(\cdot) \) represents the insured sum to be paid if the policyholder dead prematurely. In particular, [16] derived optimal strategies under minimum and maximum constraints on the insured sum; and [10] computed optimal strategies with surrender option guarantee.

In the context of financial mathematics, the work of [14, 15] has been extended by considering different models for the security prices and by assuming less restrictive utility functions. For example, with security prices modeled as semi-martingales, [19] decomposed the problem into two sub-problems and used a martingale technique to solve them. For their part, [9] and [2] used independently a martingale technique to show how to decompose the relevant Hamilton–Jacobi–Bellman equation for the problem into linear partial differential equations. Despite the existence of these theoretical results, authors such as [12] have pointed out some disadvantages of this approach. First, to obtain numerical solutions, it is often necessary to use numerical methods to solve partial differential equations; and these techniques are difficult to implement when the dimension of interesting variables increases. Also, the classical approach to attack the problem of optimal consumption and investment has been unable to link with empirical data. This lack of agreement has been documented under the name of several puzzles, namely, the equity premium puzzle (see, e.g., [13]), the risk-free rate puzzle (see, e.g., [22]), and the risk aversion puzzle (see, e.g., [8]). Keeping this in mind, Londoño proposed an alternative approach to utilities in state-complete markets. In Londoño’s framework, the consumer is supposed to have utilities that reflect his/her preference for future flows of money valued by the market when he/she is making his/her decisions (state-dependent). Another way to look at this is that people tend to value things according to their social and economic context, instead of just looking at quantitative values. For instance, people
tend to appreciate more the ability to have enough money to pay off their debts in depression times than the ability to buy luxuries in good times. In this context, the individual’s objective is to maximize

\[ E \left( \int_0^T U_1(s, H(s), c(s)) \, dt + U_3(H(T), X(T)) \right), \]

where \( H(\cdot) \) is the state price density. By considering the valuation and arbitrage theory presented in [11], Londoño developed a martingale methodology to obtain complete solutions of the problem in a quite general setting.

Hence, thinking of the drawbacks of state-independent utility models, we propose facing the problem of optimal consumption, investment, and life insurance purchase studied by [16] and [10] under the alternative approach of state-dependent utilities; following [12]. As the martingale methodology developed in [2, 1], and [9] is adaptable here; we obtain full solutions of the optimal choice problem (with and without minimum-maximum restrictions on the life insurance acquisition) in a general setup including several of the functional forms for utilities in the literature.

References


Optimal consumption, investment, and life insurance purchase


State-dependent utility with mortality risk

JAIME A. LONDOÑO
UNIVERSIDAD NACIONAL, MANIZALES, COLOMBIA
EMAIL: jaime.a.londono@gmail.com

Abstract. We solve the problem of optimal consumption and investment for intermediate consumption and final wealth where agents are assumed to have a state-dependent utility, mortality risk and have to their disposal life insurance. Life insurance is assumed to serve as collateral on depth in case of premature death, in a way that guarantees to siblings a positive inheritance. The optimization criteria are based on a state-dependent utility functional as proposed in Londoño (2009). Full solutions of the optimal consumption and portfolio problem of a consumer are obtained in a very general setting that includes incomplete markets as long as the income is hedgeable.
Minisymposium 4: Local volatility models and inverse problems

Organized by

Jorge P. Zubelli
A splitting strategy for the calibration of jump-diffusion models

Jorge P. Zubelli
IMPA, Brazil
Email: zubelli@gmail.com

Vinícius Albani
Universidade Federal de Santa Catarina, Brazil
Email: v.albani@ufsc.br

Abstract. In this talk we consider a jump-diffusion driven asset with time and price dependent volatility and vanilla option prices on such asset. We propose a splitting strategy to identify simultaneously the local-volatility surface and the jump-size distribution from quoted European prices. Our approach uses a forward Dupire-type partial-integro-differential equations for the option prices to produce a parameter-to-solution map. The ill-posed inverse problem for such map is then solved by means of a Tikhonov-type convex regularization. We shall discuss the convergence and stability of the proposed algorithm together with numerical examples that substantiate the robustness of the method both for synthetic and real data.
A New View at Local Volatility and Local Correlation

BRUNO DUPIRE
BLOOMBERG L.P., 731 LEXINGTON AVE, NEW YORK, NY 10022, USA
EMAIL: bdupire@bloomberg.net

Abstract. We show how to derive the Black-Scholes equation and the local volatility model forward equation without making use of the arguments of hedging, absence of arbitrage, change of measure, Tanaka formula... We show how we can apply the Functional Itô Calculus to obtain a fine analysis of the volatility risk. Then we look at the case of two assets and give examples of families of options that are sufficient to span any payoff and to derive the risk neutral joint density. If we assume a classical local volatility model for each of the assets we show how to compute a single local correlation function that calibrates to all two-asset European options.
Abstract. We devise a neural network parameterization of the daily snapshot of vanilla option prices, which, for a suitable architecture and nonnegative weights, is shown to be increasing in the maturity and convex in the (discounted) strike. Building on this parameterization, the local volatility calibration problem can be cast as a constrained neural network training problem, i.e. training the network to market prices of vanilla option prices under nonnegative weights condition, possibly subject to additional regularization constraints, such as imposing a priori bounds on the ensuing local volatility. The resulting local volatility deep calibration procedure is benchmarked numerically to the classical Tikhonov regularization approach.
Inversion of convex ordering: local volatility does not maximize the price of VIX futures

**Julien Guyon**
Bloomberg L.P., Quantitative Research, 731 Lexington Ave, New York, NY 10022, USA
Email: jguyon2@bloomberg.net
URL: http://cermics.enpc.fr/~guyon/

**Abstract.** It has often been stated that, within the class of continuous stochastic volatility models calibrated to vanillas, the price of a VIX future is maximized by the Dupire local volatility model. In this talk we prove that this statement is incorrect: we build a continuous stochastic volatility model in which a VIX future is strictly more expensive than in its associated local volatility model. More generally, in this model, strictly convex payoffs on a squared VIX are strictly cheaper than in the associated local volatility model. This corresponds to an inversion of convex ordering between local and stochastic variances, when moving from instantaneous variances to squared VIX, as convex payoffs on instantaneous variances are always cheaper in the local volatility model. We thus prove that this inversion of convex ordering, which is observed in the SPX market for short VIX maturities, can be produced by a continuous stochastic volatility model.
Minisymposium 5:
Machine learning methods in computational finance

Organized by
Anastasia Borovykh and Cornelis W. Oosterlee
Machine learning for quantitative finance: fast derivative pricing

SOFIE REYNERS
DEPARTMENT OF MATHEMATICS, KU LEUVEN, BELGIUM.

Abstract. In the derivatives world, daily zillion computations need to be done. Since financial models and instruments have become more and more complex, this is not always trivial and one often has to rely on time-consuming techniques. We show how machine learning algorithms can be used in this context. For many classical computations, we arrive to speed-ups of several orders of magnitude by deploying Gaussian process regression models. The price we have to pay for this extra speed is some loss of accuracy. However, we show that this reduced accuracy is often well within reasonable limits and hence very acceptable from a practical point of view.

In this talk, we focus on speeding up derivative pricing methods. In particular, we price European vanilla options, American options and exotic options under advanced market models like Heston’s model and the Variance Gamma model.

Keywords: machine learning, Gaussian processes, derivative pricing, computation time.
Reservoir computing with stochastic inputs: volatility forecasting

Lukas Gonon *
University of St. Gallen, Bodanstrasse 6, 9000 St. Gallen, Switzerland
Email: lukas.gonon@unisg.ch

Abstract. We study dynamic machine learning for discrete-time stochastic processes by means of reservoir computing, putting particular emphasis on echo state networks. Recent theoretical results prove that these systems possess universal approximation capabilities and provide error bounds for learning tasks based on them. Motivated by these results, we use reservoir computing systems to forecast realized covariances of financial time series and compare the empirical performance to state-of-the-art methods. We show that these systems demonstrate superior predictive ability and we also discuss implications on option pricing.

*The talk is based on joint works with Juan-Pablo Ortega and Lyudmila Grigoryeva
Maximum entropy approach for nonstationary time series analysis with application to volatility modelling

GANNA MARCHENKO  
UNIVERSITÀ DELLA SVIZZERA ITALIANA, LUGANO, SWITZERLAND  
EMAIL: marchenko.anna@gmail.com

Abstract. Traditional models for modelling time-dependent volatility impose explicit parametric assumptions on the data-generating process. The proposed nonparametric approach extends the maximum entropy principle to nonstationary setting and models the volatility as a persistent locally-stationary signal. Described numerical framework allows for the simultaneous identification of the optimal number of hidden regimes in data and their nonparametric regime-switching dynamics. The TV regularization is applied to recover a persistent regime-switching path and $l_1$ regularization is used for identification of the optimal number of maximum entropy distribution parameters in each of the regimes. It is demonstrated that the proposed model outperforms GARCH-type conditional volatility models with respect to the log-likelihood, information criteria, and the ability to reproduce the long memory patterns observed in the absolute returns for all considered data sets.
Overcoming the course of dimensionality with DNNs: Theoretical approximation results for PDEs

PHILIPPE VON WURSTEMBERGER
ETH ZURICH, SWITZERLAND
EMAIL: philippe.vonwurstemberger@sam.math.ethz.ch

Abstract. Artificial neural networks (ANNs) have very successfully been used in numerical simulations for a series of computational problems ranging from image classification to numerical approximations of partial differential equations (PDEs). Such numerical simulations suggest that ANNs have the capacity to very efficiently approximate high-dimensional functions and, especially, such numerical simulations indicate that ANNs seem to admit the fundamental power to overcome the curse of dimensionality when approximating the high-dimensional functions appearing in the above named computational problems. Although there are numerous results on approximation capacities of ANNs such as, e.g., the universal approximation theorem, most of them cannot explain the empirical success of ANNs when approximating high-dimensional functions. In this talk I will explain recent theoretical developments which demonstrate that ANNs can efficiently approximate solutions of high-dimensional PDEs. More precisely, I will present results revealing that the minimal required number of parameters of an ANN to approximate solutions of certain PDEs grows at most polynomially in both the reciprocal $1/\varepsilon$ of the prescribed approximation accuracy $\varepsilon > 0$ and the PDE dimension $d \in \mathbb{N}$. Those statements prove that ANNs do indeed have the capacity to overcome the curse of dimensionality in the numerical approximation of PDEs.

References


Minisymposium 6: Modelling and valuation techniques for energy markets

Organized by
Michael Coulon
Pricing swing options in electricity markets with two stochastic factors: PIDE modeling and numerical solution

MARÍA DEL CARMEN CALVO-GARRIDO*
UNIVERSIDADE DA CORUÑA, CAMPUS ELVIÑA S/N 15071 A CORUÑA, SPAIN
Email: carmen.calvo.garrido@udc.es

MATTHIAS EHRRARDT
BERGISCHEN UNIVERSITÄT WUPPERTAL, GAUSS-STRASSE 20, 42119 WUPPERTAL, GERMANY
Email: ehrhardt@math.uni-wuppertal.de

CARLOS VÁZQUEZ
UNIVERSIDADE DA CORUÑA, CAMPUS ELVIÑA S/N 15071 A CORUÑA, SPAIN
Email: carlosv@udc.es

Abstract. After the liberalization of electricity markets, prices are stated from the supply and demand principle, thus increasing volatility and uncertainty. So, companies use contracts to protect themselves from high prices and take advantage from low prices. Swing options represent a standard contract and mainly consist of path dependent options giving the holder the possibility to exercise a right multiple times over a period, where two consecutive exercises are separated by a refracting period. As it is mentioned in [3], the consideration of this refracting period avoids the exercise of all the rights at once. Following [4], in our work [2] we consider that the electricity price dynamics is described by two stochastic factors with the possibility of incorporating spikes. Thus, according to [7], the swing option pricing problem can be posed as a sequence of free boundary problems, one for each exercise right. In the presence of jumps in the electricity price, each problem is associated with a partial integro-differential equation (PIDE). Since the obstacle function associated with the free boundary problem involves the contract value with one exercise right less, additionally an initial boundary value problem linked to the same integro-differential operator has to be solved.

In order to obtain a numerical solution of the PIDE problem, we propose a Lagrange-Galerkin method for time and space discretization [1], combined with an Augmented Lagrangian Active Set (ALAS) algorithm [5] and with the treatment of the integral term by means of the explicit scheme for time discretization proposed in [6] which involves for each time step, the solution of the PIDE in the two previous steps. Moreover, we apply to the localized bounded domain artificial boundary conditions. Finally, some numerical results are presented in order to illustrate the performance of the numerical schemes and the observed properties of the solution.

References


*This work has been partially funded by MINECO of Spain (Project MTM2016-76497-R), Xunta de Galicia grants GRC2014/044 and ED431C 2018/33, including FEDER funds, and Bilateral German-Spanish DAAD Project No 57049700 HiPeCa: High Performance Calibration and Computation in Finance, Programme Acciones Conjuntas Hispano-Alemanas funded by German DAAD and the Fundación Universidad.


Wind park valuation and risk management in German intraday power markets

MICHAEL COULON
UNIVERSITY OF SUSSEX, UNITED KINGDOM
EMAIL: m.coulon@sussex.ac.uk

JONAS STRÖJBY
ALPIQ

Abstract. The rapid growth of renewables in Germany in the last decade has led to various new modeling challenges for many energy firms. Wind park owners and operators in particular require valuation and risk analysis techniques which capture the high volatility and intermittency of wind power generation, the dynamics of intraday prices and their correlation with changes in wind forecast levels. Under typical contract terms, owners of wind parks receive production volume times the spot price minus a premium $p$, while managers receive revenues dependent on how they nominate the power and rebalance their positions in the day-ahead and intraday markets. Here we present a trading and hedging strategy for determining a fair premium $p$, which can vary significantly across wind parks, for example due to their forecast variability and whether they are more or less correlated with overall wind in Germany, which drives market prices. This valuation problem is of significant interest to many market participants, including investors and policy makers looking to further grow the penetration of renewables.
Efficient gas swing and storage valuation with neural network and PDE/MC

KLAUS SPANDEREN
UNIPER
EMAIL: klaus.spanderen@uniper.energy

Abstract. Swing and storage contracts are fundamental tools for supply and risk management of gas markets. These options are usually valued by means of approximation formulas, partial differential equations or least square Monte-Carlo methods. Typically, the valuation of swings and storages is complicated and computational demanding. We trained a deep neural network on a large set of market reflective market data to obtain real-time valuation and compare it with state of the art stochastic local volatility models. In our experiments we show that the valuation error of the net is of the same order as the LSMC error of the training data, however the computation time is several magnitudes faster.
Forecasting intra-day price spread density matrices for electricity storage operations

Derek Bunn
London Business School
Email: dbunn@london.edu

Ekaterina Abramova
London Business School

Abstract. This paper formulates dynamic density functions, based upon skewed-t and similar representations, to model and forecast electricity price spreads as a matrix between different hours of the day. This supports an optimal day ahead storage and discharge schedule, and thereby facilitates a bidding strategy for a merchant arbitrage facility into the day-ahead auctions for wholesale electricity. The four latent moments of the density functions are dynamic and conditional upon exogenous drivers, thereby permitting the mean, variance, skewness and kurtosis of the densities to respond hourly to such factors as weather and demand forecasts. The best specification for each spread is selected based on the Pinball Loss function, following the closed form analytical solutions of the cumulative density functions. Those analytical properties also allow the calculation of risk associated with the spread arbitrages. From these spread densities, the optimal daily operation of a battery storage facility is determined, if it seeks revenue from arbitrage. Some observations on multiple, alternative revenue streams for battery operations are advanced.
Minisymposium 7: Numerical Methods for PDEs in finance

Organized by
Matthias Ehrhardt and Jan ter Maten
Operator splitting schemes for the two-asset Merton jump-diffusion model

LYNN BOEN*
MIDDELHEIMAAN 1, B-2020 ANTWERP, BELGIUM
EMAIL: lynn.boen@uantwerpen.be

KAREL J. IN 'T HOUT*
MIDDELHEIMAAN 1, B-2020 ANTWERP, BELGIUM
EMAIL: karel.inthout@uantwerpen.be

Abstract. Under the two-asset Merton jump-diffusion model, the value of a European-style option satisfies a two-dimensional time-dependent partial integro-differential equation (PIDE). We study seven recent and novel operator splitting schemes when applied to this PIDE, with a keen focus on implicit-explicit (IMEX) and alternating direction implicit (ADI) methods. Each of the schemes conveniently treats the nonlocal integral part in an explicit fashion. Through ample numerical experiments we investigate the convergence behaviour of the different splitting schemes and study their relative performance.

Under the two-asset Merton jump-diffusion model the value $v = v(s_1, s_2, t)$ of a European-style option, with maturity date $T > 0$ and $s_i$ ($i = 1, 2$) the price of asset $i$ at time $T - t$, satisfies the PIDE

$$
\frac{\partial v}{\partial t} = \frac{1}{2} \sigma_1^2 s_1^2 \frac{\partial^2 v}{\partial s_1^2} + \rho \sigma_1 \sigma_2 s_1 s_2 \frac{\partial^2 v}{\partial s_1 \partial s_2} + \frac{1}{2} \sigma_2^2 s_2^2 \frac{\partial^2 v}{\partial s_2^2} + (r - \lambda \kappa_1) s_1 \frac{\partial v}{\partial s_1} + (r - \lambda \kappa_2) s_2 \frac{\partial v}{\partial s_2} - (r + \lambda) v + \lambda \int_0^\infty \int_0^\infty v(s_1 y_1, s_2 y_2, t) f(y_1, y_2) dy_1 dy_2
$$

for $s_1 > 0$, $s_2 > 0$, $0 < t \leq T$. Here $f$ is the probability density function of a bivariate lognormal distribution. After semidiscretization of the two-dimensional Merton PIDE, a large system of ordinary differential equations (ODEs) is obtained,

$$
V'(t) = AV(t)
$$

for $0 < t \leq T$. The given matrix $A$ can be decomposed as $A = A^{(M)} + A_1 + A_2 + A^{(J)}$, where $A^{(M)}$ represents the mixed derivative term, $A_i$ ($i = 1, 2$) represents all spatial derivatives in the $s_i$-direction, and $A^{(J)}$ represents the integral part. Since the integral part is nonlocal, the matrix $A^{(J)}$ is dense and this poses a particular challenge for an efficient implicit temporal discretization. In our talk we study seven recent and novel operator splitting schemes of the IMEX and ADI kind, where the integral part is always conveniently treated in an explicit fashion. The convergence behaviour and the relative performance of these schemes shall be investigated in ample numerical experiments for European put-on-the-min and put-on-the-average options [1, 2].

References


A penalty-like method for CVA pricing by a PDE model

CHRISTINA CHRISTARA∗
DEPART. OF COMPUTER SCIENCE, UNIVERSITY OF TORONTO
EMAIL: ccc@cs.toronto.edu
URL: http://www.cs.toronto.edu/~ccc

Yuwei Chen†
DEPART. OF COMPUTER SCIENCE, UNIVERSITY OF TORONTO
EMAIL: ywchen@cs.toronto.edu

Abstract. Counterparty risk is the risk to each party of a contract that a counterparty may not live up to its contractual obligations. Counterparty risk must be evaluated properly, and the risk neutral value of a derivative must be adjusted accordingly. For the pricing of the credit valuation adjustment (CVA), or, more generally, the total valuation adjustment, known as XVA, we adopt a Black-Scholes PDE model with additional nonlinear terms. For the discretization, we use standard second order differences in space and Crank-Nicolson timestepping. For the treatment of the nonlinearity, we formulate a penalty-like iteration. We present numerical experiments indicating that the penalty method converges in about one iteration per timestep, irrespectively of the discretization size, and that second order convergence is exhibited.

1. Introduction. After the 2007-2008 financial crisis, the accurate evaluation of counterparty risk, also known as default risk [5, 6], and the appropriate adjustment of a financial derivative’s risk neutral value have become increasingly important. This adjustment is called credit valuation adjustment (CVA). There are several types of adjustments, the sum of which is often referred to by the generic name XVA. We are interested in calculating the the total valuation adjustment to the price of a derivative, that is, calculate

\[ U = \hat{V} - V \]

where \( \hat{V} \) the value taking into account credit risk, \( V \) the value without credit risk, and \( U \) the XVA. According to the International Swaps and Derivatives Association (ISDA) 2002 Master Agreement, in case either party defaults, the value of the derivative is determined by a mark-to-market rule \( M \), which can be either \( \hat{V} \) or \( V \). By convention, the positive derivative values correspond to seller’s assets and counterparty liabilities, while the negative derivative values correspond to seller’s liabilities and counterparty assets. The positive and negative values of any security \( W \) are denoted by \( W^+ \) and \( W^- \), respectively. Note that \( W^+ = \max\{W, 0\} \), \( W^- = \min\{W, 0\} \) and \( W = W^+ + W^- = \max\{W, 0\} + \min\{W, 0\} \).

In this paper, we focus on a PDE model for \( U \), which is nonlinear in the case \( M = \hat{V} \). In Section 2, we define some notations and introduce the nonlinear PDE model. In Section 3, we consider numerical methods for approximating the solution of the PDE. In particular, we focus on a penalty-like method to handle the nonlinearity. In Section 4, we present numerical results that demonstrate the performance of the method. We conclude in Section 5.

2. PDE model to price bilateral counterparty risk.

2.1. Notation. Let \( S \) be the underlying stock value, \( \sigma \) the volatility in \( S \), \( r \) the risk-free interest rate, \( \gamma \) the dividend yield of \( S \), \( q \) the stock repo rate, \( r_B \) the seller’s bond yield, \( r_C \) the counterparty’s bond yield, \( \lambda_B = r_B - r \) (bank hazard rate, instantaneous probability of bank default), \( \lambda_C = r_C - r \) (counterparty hazard rate), \( R_B \) the recovery percentage on \( M \) if seller defaults, \( R_C \) the recovery percentage on \( M \) if counterparty defaults, \( r_F \) the seller’s funding rate for borrowed cash, where \( r_F = r \) if derivative can be used as collateral, and \( r_F = r + (1 - R_B)\lambda_B \) if derivative cannot be used as collateral, \( s_F = r_F - r \).

∗This research was partly supported by the Natural Sciences and Engineering Research Council of Canada (NSERC)
†This research was partly supported by the Ontario Graduate Scholarship (OGS)
Define also the differential operator

\[ A \equiv \sigma^2 S^2 \frac{\partial^2}{\partial S^2} + (q - \gamma)S \frac{\partial}{\partial S}. \]

2.2. The PDE problem. In [3], using replicating portfolio arguments, it is shown that, under certain assumptions, \( \hat{V} \) satisfies the PDE

\[ \frac{\partial \hat{V}}{\partial t} + A \hat{V} - r \hat{V} = (\lambda_B + \lambda_C)\hat{V} - \lambda_B (R_B M^- + M^+) - \lambda_C (R_C M^+ + M^-) + s_F M^+ \]

with initial conditions

\[ \hat{V}(T, S) = H(S), \]

where \( t \) is the (forward) time variable and \( T \) the maturity time of the derivative.

Note also that \( V \) satisfies the Black-Scholes PDE

\[ \frac{\partial V}{\partial t} + AV - r V = 0 \]

with initial conditions

\[ V(T, S) = H(S). \]

When \( M = \hat{V} \), and recalling that \( \hat{V} = \hat{V}^+ + \hat{V}^- \), PDE (1), becomes

\[ \frac{\partial \hat{V}}{\partial t} + A \hat{V} - r \hat{V} = \lambda_B (1 - R_B) \hat{V}^- + \lambda_C (1 - R_C) \hat{V}^+ + s_F \hat{V}^+ \]

with initial conditions

\[ \hat{V}(T, S) = H(S). \]

Using \( \hat{V} = V + U \) and PDE (2), PDE (3) can be written for \( U \) as

\[ \frac{\partial U}{\partial t} + AU - rU = \lambda_B (1 - R_B) (V + U)^- + \lambda_C (1 - R_C) (V + U)^+ + s_F (V + U)^+ \]

with initial conditions

\[ U(T, S) = 0. \]

Note that, in (4), \( V \) is considered known. It can be computed by the Black-Scholes formula, or numerically. Note also that \( \hat{V} \) in (3) and \( U \) in (4) are unknown, and, since these PDEs involve unknowns in the max or min notation, they are nonlinear.

3. Numerical solution of nonlinear PDE. In [1, 2], a numerical method to approximate the solution of (4) is proposed and tested. The method involves using a characteristics (semi-Lagrangian) method for time-stepping, a finite element method for the space discretization, and a fixed-point iteration scheme to handle the nonlinearity. The authors of [1, 2] report experimental results indicating that their method is first order convergent.

In this paper, we focus on the iteration scheme for the nonlinearity. In order to describe our proposed iteration scheme, we present a simplified version of the iteration scheme in [1, 2]. For this reason, we assume that the time-stepping is handled by the \( \theta \)-method, which, for \( \theta = \frac{1}{2} \) and \( \theta = 1 \) becomes the Crank-Nicolson and Backward Euler methods, respectively. We also assume that the space discretization is handled by a standard (second-order) finite difference method. With

\[ f(U, V) \equiv - (\lambda_B (1 - R_B) (V + U)^- + \lambda_C (1 - R_C) (V + U)^+ + s_F (V + U)^+) \]

and with \( \tau = T - t \), the PDE (4) is written in a more concise and more convenient for numerical computation form as

\[ \frac{\partial U}{\partial \tau} = AU - rU + f(U, V). \]
Let $\tau_j, j = 0, \ldots, N_t$, be the timesteps at which the solution is computed, with $\tau_0 = 0 < \tau_1 < \ldots < \tau_{N_t} = T$, and $\Delta \tau^j = \tau_j - \tau_{j-1}$ be the jth time stepsize. Let $u_j, j = 0, \ldots, N_t$, denote the computed solution vector arising from approximate values of $U$ at the spatial gridpoints, and at timestep $j$, with $u^0$ being the initial condition vector. Let $u^{j,k}, k = 0, \ldots, \text{maxit}$, denote the computed solution vector at iteration $k$ of timestep $j$, with maxit the maximum number of iterations allowed. Let $v^j, j = 0, \ldots, N_t$, and $P^j, j = 0, \ldots, N_t$, denote the vectors arising from the values of $V$ and $f(U, V)$, respectively, at the spatial gridpoints, and at timestep $j$. Let also $A$ be the matrix arising from the space discretization of $A U - r U$, and $I$ be the identity matrix of same dimension. We assume, for simplicity, that the spatial gridpoints remain the same at all timesteps. Then a simplified version of the fixed-point iteration method in [1, 2] at timestep $j$ can be presented in high-level notation as follows.

**Algorithm 1:** Fixed-point iteration for (6) at step $j$, with $\theta$-timestepping
1. initialize $u^{j,0} = u^{j-1}$
2. for $k = 1, \ldots, \text{maxit}$
3. solve $(I - \theta \Delta \tau^j A) u^{j,k} = (I + (1 - \theta) \Delta \tau^j A) u^{j-1} + \theta \Delta \tau^j f(u^{j,k-1}, v^j) + (1 - \theta) \Delta \tau^j f(u^{j-1}, v^{j-1})$
4. check stopping criterion
5. endfor ($k$)
6. set $u^j = u^{j,k}$

Note that the above iteration method is considered to be explicit, in the sense that the matrix solved at each iteration is the same and only the right-hand side vector changes.

We next present another iteration method for handling the nonlinearity in (6). We refer to it as discrete penalty iteration as it is motivated by the same-name method in [4], designed to handle the nonlinear PDE arising from the linear complementarity problem in American option pricing. For this reason, with generic vectors $u$ and $v$, we define the diagonal penalty matrix $P = P(u, v)$ by

$$P_{i_1, i_2}(u, v) \equiv \begin{cases} -\lambda_R (1 - R_B) & \text{if } i_1 = i_2 \text{ and } v_{i_1} < -u_{i_1} \\ -\lambda_C (1 - R_C) - s_F & \text{if } i_1 = i_2 \text{ and } v_{i_1} \geq -u_{i_1} \\ 0 & \text{otherwise.} \end{cases}$$

(7)

Note that, with the above matrix $P$, the vector arising from the discretized form of $f(U, V)$ can be written as

$$f(u, v) = P(u, v)(u + v),$$

where, since $P(u, v)$ depends on $u$, there is nonlinearity between $P(u, v)$ and $u$. Note also that, in a typical case, $\lambda_R \geq 0, \lambda_C \geq 0$, and $s_F \geq 0$, and we have $P_{i_1, i_2}(u, v) \leq 0$.

Let now $P^k = P(u^{j,k}, v^j)$, where, for brevity, in the notation $P^k$, we omitted the superscript $j$. Using the penalty matrix $P^k$, the proposed discrete penalty iteration for (6) is as follows.

**Algorithm 2:** Discrete penalty iteration for (6) at step $j$, with $\theta$-timestepping
1. initialize $u^{j,0} = u^{j-1}$ and $P^0 = P(u^{j,0}, v^j)$
2. for $k = 1, \ldots, \text{maxit}$
3. solve $(I - \theta \Delta \tau^j (A + P^{k-1})) u^{j,k} = (I + (1 - \theta) \Delta \tau^j A) u^{j-1} + \theta \Delta \tau^j P^{k-1} v^j + (1 - \theta) \Delta \tau^j f(u^{j-1}, v^{j-1})$
4. compute $P^k = P(u^{j,k}, v^j)$
5. check stopping criterion
6. endfor ($k$)
7. set $u^j = u^{j,k}$

Note that the above iteration method is considered to be implicit, in the sense that the matrix solved at each iteration depends on the iteration index $k$. However, since the matrix solved at each iteration is adjusted by only a diagonal matrix, the sparsity structure of the matrix remains the same, which is also the same structure as of the matrix corresponding to a linear PDE with the same differential operator as in (6). Furthermore, since, in a typical case, $P_{i_1, i_1}(u, v) \leq 0$, and also $A_{i_1, i_1} < 0$, the matrix $P$ enhances the diagonal dominance of $A$.

### 3.1. Stopping criterion.
In our method, motivated by [4], we use as stopping criterion

$$P^k = P^{k-1} \text{ or } \left( \max_i \frac{|u_{i,k} - u_{i,k-1}|}{\max(1, |u_{i,k}|)} \leq \text{tol} \right).$$

(8)
4. Numerical results. We present results from pricing the XVA of a 5-year European Put option $(T = 5)$ with strike $K = 15$ and the following parameters: $\sigma = 0.25$, $r = 0.03$, $q - \gamma = 0.015$, $r_B = 0.05$, $r_C = 0.08$, $R_B = R_C = 0.4$, $r_F = r + (1 - R_B) \lambda_B$. For this problem, we have an exact solution formula as given in [3], thus we can calculate exact errors. We truncate the semi-infinite domain of $S$ to $[0, 12K]$, and discretize the spatial domain in $N$ subintervals, so that the (nonuniform) gridpoints are concentrated around the strike $K$. The number of timesteps is $N_t$, and $\Delta \tau = T/N_t$. We formulate Dirichlet boundary conditions using the techniques in [1, 2] for the left end, and linear far-field conditions as in [7].

In Table 1, we report, for several discretization sizes, the maximum in absolute value error over all gridpoints at time $t = 0$ ($\tau = T$), the associated convergence rates, the total and average per timestep number of iterations. The penalty iteration tolerance was set to $10^{-7}$. We notice that the average number of penalty-like iterations is very close to 1, irrespectively of the grid size, and note that this is optimal. Essentially, in each of the first few timesteps two or three iterations are needed, and in each of the remaining timesteps just one iteration. The convergence rate is stably close to 2.

5. Conclusions. The penalty-like method presented is very efficient in handling the nonlinearity of the XVA pricing PDE and together with standard space and time discretization techniques produces second order approximations.

References


BDF finite difference schemes for diffusion equations with obstacle

Olivier Bokanowski
Laboratoire Jacques-Louis Lions, Univ. Paris-Diderot (Paris 7), and Ensta ParisTech, France
Email: olivier.bokanowski@math.univ-paris-diderot.fr

Kristian Debrabant
University of Southern Denmark, Department of Mathematics and Computer Science,
Odense, Danmark
Email: debrabant@imada.sdu.dk

Abstract. We study Backward Differentiation Formula (BDF) 2 and BDF 3 schemes for the approximation of one-dimensional diffusion equations with an obstacle term, of the form

$$\min (v_t - a(t, x)v_{xx} + b(t, x)v_x + r(t, x)v, v - \varphi(t, x)) = f(t, x).$$

For the scheme building on BDF2, we discuss unconditional stability, present an $L^2$–error estimate and show numerically second order convergence, in both space and time, unconditionally on the ratio of the mesh steps.

Two academic problems for parabolic equations with an obstacle term with explicit solutions and the American option problem in mathematical finance are used for numerical tests.
High-order compact finite difference scheme for option pricing in stochastic volatility jump models

BERTRAM DÜRING
DEPARTMENT OF MATHEMATICS, UNIVERSITY OF SUSSEX, PEVENSEY II, BRIGHTON, BN1 9QH, UNITED KINGDOM
EMAIL: bd80@sussex.ac.uk
URL: http://users.sussex.ac.uk/~bd80

ALEXANDER PITKIN
DEPARTMENT OF MATHEMATICS, UNIVERSITY OF SUSSEX, PEVENSEY II, BRIGHTON, BN1 9QH, UNITED KINGDOM
EMAIL: A.H.Pitkin@sussex.ac.uk

Abstract. We present a new high-order compact finite difference scheme [4, 5] for option pricing in stochastic volatility jump models, e.g. in Bates model [1] and the stochastic volatility with contemporaneous jump model (SVCJ) model [2]. In such models the option price is determined as the solution of a partial integro-differential equation. The implicit-explicit scheme is based on the approaches in Düring and Fournié [3] and Salmi et al. [6]. The scheme is fourth order accurate in space and second order accurate in time. Numerical experiments for the European option pricing problem are presented. We validate the stability of the scheme numerically and compare its performance to standard finite difference and finite element methods. The new scheme outperforms a standard discretisation based on a second-order central finite difference approximation in all our experiments. At the same time, it is very efficient, requiring only one initial LU-factorisation of a sparse matrix to perform the option price valuation. Compared to finite element approaches, it is very parsimonious in terms of memory requirements and computational effort, since it achieves high-order convergence without requiring additional unknowns, unlike finite element methods with higher polynomial order basis functions. The new high-order compact scheme can also be useful to upgrade existing implementations based on standard finite differences in a straightforward manner to obtain a highly efficient option pricing code.

References
Minisymposium 8:
Qualitative and quantitative aspects of nonlinear PDEs based models arising in mathematical finance

Organized by
Daniel Ševčovič
Abstract. In this work, we derive new linear and nonlinear partial differential equations (PDEs) models for pricing American options and total value adjustment in the presence of counterparty risk. Moreover, stochastic spreads are considered, which increases the dimension of the problem.

1. Introduction. Since the last crisis, suitable relevant adjustments are taken into account when pricing derivatives, their goal being to mitigate the counterparty risk associated to derivatives contract. The set of these adjustments is continuously increasing and is known as total value adjustment or XVA. In [1], we have studied the XVA modelling and numerical methods based on partial differential equations (PDEs) for European and American options depending on one stochastic factor, as a result of considering constant intensities of default.

However, default intensities from counterparties do not always exhibit constant behaviour. In [3], where both risky counterparties are considered to have stochastic intensities of default, a three underlying stochastic factors model is obtained. Following this work, we extend the model introduced in [1] to the case of stochastic spreads. A similar approach has been developed in [2] for European options.

2. Mathematical model. In this section, we obtain the models for American options considering counterparty risk. We assume an investor as a risky counterparty and consider that the issuer’s intensity of default is null. Thus, the dynamics of underlying asset price $S_t$ and the short term CDS spread of the investor $h_t$ are modelled by the following system of stochastic differential equations:

$$dS_t = (r(t) - q(t)) S_t dt + \sigma^S(t) S_t dW^S_t,$$

$$dh_t = (\mu^h(t) - M^h(t) \sigma^h(t)) dt + \sigma^h(t) dW^h_t,$$

where $(r(t) - q(t))$ and $(\mu^h(t) - M^h(t) \sigma^h(t))$ are the (respectively) drifts of the processes. Moreover, $r(t)$ denotes the risk-free interest rate, $q(t)$ is the asset dividend yield rate, $M^h(t)$ is the market price of investor’s credit risk, $\sigma^S(t)$ and $\sigma^h(t)$ are the volatility functions, and $W^S_t$ and $W^h_t$ are two correlated Wiener processes (i.e., $\rho dt = dW^S_t dW^h_t$) such that $\rho$ is the instantaneous correlation between $S_t$ and $h_t$.

In order to derive the value of American options with counterparty risk, we consider the self-financing portfolio $\Pi_t$, which at time $t$ is given by

$$\Pi_t = \alpha(t) H(t) + \beta(t) + \gamma(t) \text{CDS}(t,T) + \varepsilon(t) \text{CDS}(t, t + dt) + \Omega(t) B(t, t + dt).$$
Moreover, with the aim to avoid arbitrage opportunities, we introduce the following hedging inequality:
\[ d\hat{V}_t \leq d\Pi_t. \]  
Next, applying Itô’s Lemma for jump diffusion processes, and replacing the change of the portfolio and the change of the derivative in the hedging inequality (4), we deduce a complementarity problem modelling the American option value with counterparty risk:
\[
\begin{aligned}
&\mathcal{L}(\hat{V}) = \frac{\partial \hat{V}}{\partial t} + \mathcal{L}_{sh} \hat{V} + \frac{\Delta \hat{V}}{1-R} h - f\hat{V} \leq 0 \\
&\hat{V}(t, S, h) \geq G(S) \\
&\mathcal{L}(\hat{V})(\hat{V} - G) = 0 \\
&\hat{V}(T, S, h) = G(S),
\end{aligned}
\]  
where \( G(S) \) represents the option payoff, \( \Delta \hat{V} = \hat{V}_M + \hat{V} - \hat{V} \) and \( M \) is the mark-to-market close-out.

Moreover, the differential operator \( \mathcal{L}_{sh} \) is
\[
\mathcal{L}_{sh} \equiv \frac{1}{2}(\sigma S)^2 \frac{\partial^2 V}{\partial S^2} + \frac{1}{2}(\sigma h)^2 \frac{\partial^2 V}{\partial h^2} + \rho \sigma S \frac{\partial^2 V}{\partial h \partial S} + \frac{(r - q)}{1-R} h \frac{\partial V}{\partial h}.
\]

We consider two possible values for the mark-to-market close-out, which leads to a nonlinear or linear problem:

- If \( M = \hat{V} \), we deduce the following nonlinear complementarity problem:
\[
\begin{aligned}
&\mathcal{L}_1(\hat{V}) = \frac{\partial \hat{V}}{\partial t} + \mathcal{L}_{sh} \hat{V} - f\hat{V} - h\hat{V}^+ \leq 0, \quad \text{in } [0, T) \times (0, \infty) \times (0, \infty) \\
&\hat{V}(t, S, h) \geq G(S) \\
&\mathcal{L}_1(\hat{V})(\hat{V} - G) = 0 \\
&\hat{V}(T, S, h) = G(S).
\end{aligned}
\]  

- If \( M = V \), the following linear complementarity problem is derived:
\[
\begin{aligned}
&\mathcal{L}_2(\hat{V}) = \frac{\partial \hat{V}}{\partial t} + \mathcal{L}_{sh} \hat{V} - \left( \frac{h}{1-R} + f \right) \hat{V} \\
&-((1-R)V^+ - V)h \frac{h}{1-R} \leq 0, \quad \text{in } [0, T) \times (0, \infty) \times (0, \infty) \\
&\hat{V}(t, S, h) \geq G(S) \\
&\mathcal{L}_2(\hat{V})(\hat{V} - G) = 0 \\
&\hat{V}(T, S, h) = G(S).
\end{aligned}
\]  

Moreover, the calculus of the XVA value, \( U = \hat{V} - V \), requires the previous computing of the counterparty risk-free American option value, which is modelled by the classical Black-Scholes equation for American options.

In order to solve model (5) by a finite element method, we first localize the problem on a bounded domain. For this purpose, let us consider \( \Omega = (0, S_\infty) \times (0, h_\infty) \) for large enough values of \( S_\infty \) and \( h_\infty \), so that their choice does not affect the solution in the domain of financial interest. We need to impose appropriate boundary conditions on the risky derivative value problem in the bounded domain. We consider the same boundary conditions than for \( V \) and \( \hat{V} \) as in the case of European options [2]. Then, focusing on the nonlinear case, we deduce the following nonlinear complementarity problem for
the XVA:
\[
\begin{align*}
\mathcal{L}_t(U) &= \frac{\partial U}{\partial t} + \mathcal{L}_{sh}U - fU - h(U + V)^+ \leq -\frac{\partial V}{\partial t} - \mathcal{L}_S V + fV, \quad t \in [0,T), \quad (S,h) \in \Omega \\
U(t,S,h) &\geq G(S) - V(t,S) \\
\left[ \mathcal{L}_t(U) - \left( -\frac{\partial V}{\partial t} - \mathcal{L}_S V + fV \right) \right] [U - (G(S) - V(t,S))] &= 0 \\
U(T,S,h) &= 0 \\
U(t,0,h) &= 0 \\
U(t,S_{\infty},h) &= 0 \\
U(t,S,0) &= 0 \\
(\mathcal{A}U \cdot \vec{n})(\tau,S,h_{\infty}) &= 0.
\end{align*}
\]

3. Mathematical analysis. In this section, we prove the existence and uniqueness of solution for the XVA problem (8) for a given function $V$. Then, taking into account the existence and uniqueness of solution $V$ for the classical Black-Scholes problem, we constate the existence and uniqueness of solution for (6). Introducing the time to maturity variable, $\tau = T - t$, as well as the new variables and unknown
\[
x = \ln S, \quad u(\tau,x,h) = U(t,S,h), \quad v(\tau,x) = V(t,S),
\]
we pose the nonlinear complementarity problem (8) in $(0,T) \times \tilde{\Omega}$ as follows:
\[
\begin{align*}
\mathcal{L}_\tau(u) &= \frac{\partial u}{\partial \tau} + Au - \Phi(\tau,u) \geq \ell, \quad (x,h) \in \tilde{\Omega}, \quad \tau \in (0,T) \\
u &\geq \psi \\
[\mathcal{L}_\tau(u) - \ell][u - \psi] &= 0 \\
u(0,S,h) &= 0 \\
u(\tau,x_0,h) &= 0 \\
u(\tau,x_{\infty},h) &= 0 \\
u(\tau,x,0) &= 0 \\
(\mathcal{A}v \cdot \vec{n})(\tau,x,h_{\infty}) &= 0.
\end{align*}
\]

**Theorem 3.1.** The following statements are satisfied:

1. The continuous operator $A$ satisfies Garding’s inequality, i.e.:
\[
(Az,z) \geq \omega_1 \|z\|^2_{H_0^1(\tilde{\Omega})} - \omega_2 \|z\|^2_{L^2(\tilde{\Omega})} \quad \forall z \in H_0^1(\tilde{\Omega}),
\]
for $\omega_1 > 0$ and $\omega_2 \in \mathbb{R}$.

2. $\ell \in L^2(0,T;L^2(\tilde{\Omega})) \subset L^2(0,T;W^*)$.

3. Let $D(\phi) = \left\{ z \in H_0^1(\tilde{\Omega}) / (\phi(z) < \infty) \right\}$ and $u_0 = u(0,x,h)$. Then, $u_0 \in \overline{D(\phi)}$.

4. $\Phi(\tau,\varphi)$ is Lipschitz continuous on variable $\varphi$, i.e.
\[
\|\Phi(\tau,\varphi_1) - \Phi(\tau,\varphi_2)\|_{L^2(\tilde{\Omega})} \leq L_G \|\varphi_1 - \varphi_2\|_{H_0^1(\tilde{\Omega})}.
\]

Therefore, by [4] the nonlinear complementarity problem (9) has a unique solution $u \in L^2(0,T;H_0^1(\tilde{\Omega})) \cap C([0,T];L^2(\tilde{\Omega}))$; in particular $u \in W^{1,2}(0,T;L^2(\tilde{\Omega}))$ and satisfies
\[
\|u\|_{W^{1,2}(0,T;L^2(\tilde{\Omega}))} \leq C_1 \left(1 + \|u_0\|_{L^2(\tilde{\Omega})} + \|\ell\|_{L^2(0,T;H_0^1(\tilde{\Omega}))}\right).
\]
4. Numerical simulation. The numerical approximation is mainly based on finite elements combined with the method of characteristics. More precisely, we introduce the finite element space and convex set:

\[ W_h = \{ \varphi_h \in C(\Omega) / \tilde{\varphi}|_{T_j} \in P_1, \forall T_j \in T \}, \]
\[ K_h = \{ \varphi_h \in W_h / \varphi_h = \hat{V} \text{ on } \Gamma^+_1 \cup \Gamma^+_2 \text{ and } \varphi_h \geq G(S) \}, \]

in order to find \( \hat{V}^{n+1,k+1}_h \in K_h \) satisfying the boundary conditions and such that:

\[
\int_{\Omega} (1 + \Delta \tau^n f) \hat{V}^{n+1,k+1}_h (\varphi_h - \hat{V}^{n+1,k+1}_h) \, dS \, dh + \Delta \tau^n \int_{\Omega} A \nabla \hat{V}^{n+1,k+1}_h \nabla (\varphi_h - \hat{V}^{n+1,k+1}_h) \, dS \, dh \\
\geq \int_{\Omega} (\hat{V}^{n}_h \circ \chi^n)(\varphi_h - \hat{V}^{n+1,k+1}_h) \, dS \, dh - \Delta \tau^n \int_{\Omega} h(\hat{V}^{n+1,k}_h)^+ (\varphi_h - \hat{V}^{n+1,k+1}_h) \, dS \, dh,
\]

for all \( \varphi_h \in K_h \).

The previous variational problem is discretized by a \( P_1 \) Lagrange finite elements method. This technique leads to an obstacle-like system of linear equations, which is solved by an augmented Lagrangian active set (ALAS) algorithm [5]. Figure 1 shows the total value adjustment and the exercise region.

References


On solutions of a partial integro-differential Black-Scholes equation in Bessel potential spaces

Daniel Ševčovič
Faculty of Mathematics, Physics and Informatics, Comenius University, Mlynská Dolina, 84248 Bratislava, Slovak Republic
Email: sevcovic@fmph.uniba.sk

José Cruz
University of Lisbon, Portugal

Abstract. The classical linear Black-Scholes model based on a solution to a linear parabolic partial differential equation is a popular model used in financial markets. But it relies on several restrictive assumptions such as the liquid, complete and frictionless market. Relaxing the complete market hypothesis and assuming a Lévy stochastic process dynamics for the underlying stock price we obtain a nonlocal partial integro-differential equation as a model for pricing options. We consider a model for pricing vanilla call and put options on underlying assets following Lévy stochastic process. Using the theory of abstract semilinear parabolic equations we proved existence and uniqueness of solutions in the Bessel potential space. Our aim is to generalize known existence results for a wide class of Lévy measures including those having strong singular kernel. We also prove existence and uniqueness of solutions to the penalized PIDE representing approximation of the linear complementarity problem arising in pricing American style of options under Lévy stochastic process.
Dynamic portfolio optimization via a Hamilton-Jacobi-Bellman equation

Sôňa Kilianová
Email: sôna.kilianova@fmph.uniba.sk
Mária Trnovská
Daniel Ševčovič
Email: sevcovic@fmph.uniba.sk

Abstract. We consider a problem of dynamic stochastic portfolio optimization based on terminal expected utility optimization, which leads to solving a Hamilton-Jacobi-Bellman (HJB) equation. The HJB equation is transformed to a simpler quasi-linear PDE by means of the so-called Riccati transformation. An auxiliary quadratic programming problem is obtained, which uses a vector of expected asset returns and a covariance matrix of returns as input parameters. This basic concept can be extended to modified portfolio optimization problems. Since this problem can be sensitive to the input data, we modify the problem from fixed input parameters to worst-case optimization over convex or discrete uncertainty sets both for asset returns and covariance matrix. We also look at how solutions change if we allow for non-constant risk aversion parameter. Next, we can investigate the case when not only terminal expected utility is to be maximized, but also overall inter-temporal utility during entire investment period. Qualitative properties of the corresponding value functions and solutions to selected types of problems are analyzed along with providing illustrative numerical examples.
Option pricing in illiquid markets with jumps

JOSE M. T. S. CRUZ
ISEG, UNIVERSITY OF LISBON, RUA DE QUELHAS 6, 1200-781 LISBON, PORTUGAL
EMAIL: jmcruz@iseg.ulisboa.pt

DANIEL ŠEVČOVIČ
COMENIUS UNIVERSITY IN BRATISLAVA, MLYNSKÁ DOLINA, 84248 BRATISLAVA, SLOVAKIA
EMAIL: sevcovic@fmph.uniba.sk

Abstract. The classical linear Black–Scholes model for pricing derivative securities is a popular model in financial industry. It relies on several restrictive assumptions such as completeness, and frictionless of the market as well as the assumption on the underlying asset price dynamics following a geometric Brownian motion. The main purpose of this paper is to generalize the classical Black–Scholes model for pricing derivative securities by taking into account feedback effects due to an influence of a large trader on the underlying asset price dynamics exhibiting random jumps. The assumption that an investor can trade large amounts of assets without affecting the underlying asset price itself is usually not satisfied, especially in illiquid markets. We generalize the Frey–Stremme nonlinear option pricing model for the case the underlying asset follows a Lévy stochastic process with jumps. We derive and analyze a fully nonlinear parabolic partial-integro differential equation for the price of the option contract. We propose a semi-implicit numerical discretization scheme and perform various numerical experiments showing influence of a large trader and intensity of jumps on the option price.

Keywords: Nonlinear partial integro-differential equation, Lévy measure, Finite difference approximation
Pension saving strategies based on Samuelson’s lifecycle theory

IGOR MELICHERCÍK
FACULTY OF MATHEMATICS, PHYSICS AND INFORMATICS, COMENIUS UNIVERSITY, MLÝNSKÁ DOLINA, 84248 BRATISLAVA, SLOVAK REPUBLIC
EMAIL: igor.melichercik@fmph.uniba.sk

ALEŠ ČERNÝ
CASS BUSINESS SCHOOL, CITY UNIVERSITY OF LONDON, UNITED KINGDOM

GÁBOR SZŰCS
FACULTY OF MATHEMATICS, PHYSICS AND INFORMATICS, COMENIUS UNIVERSITY, MLÝNSKÁ DOLINA, 84248 BRATISLAVA, SLOVAK REPUBLIC

Abstract. We consider optimal investment for an individual pension savings plan in receipt of gradual contributions against which one cannot borrow, using expected power utility as the optimality criterion. Risky assets are modelled using a multidimensional Ito process. In addition, there is a risk-free asset with deterministic value process. It is known that, in the case of no credit constraints, the optimal investment can be described by constant proportions of assets out of lifetime wealth (including current savings and present value of future contributions). On the other hand, in the presence of credit constraints, this solution is no longer applicable, and the optimal investment leads to so-called stochastic lifestyling, whereby for low levels of savings it is optimal to invest fully in stocks and then gradually switch to safer assets as the level of savings increases. We highlight that not only does the leverage between risky and safe assets change, but also mutual proportions of investments in different risky assets vary over time.

We formulate a problem of dynamic stochastic programming with credit constraints and present the corresponding Hamilton-Jacobi-Bellman (HJB) equation. The fully nonlinear HJB equation is transformed (using Riccati transformation) into equivalent quasi-linear parabolic equation for which we prove (under natural conditions) the existence and uniqueness of sufficiently smooth solution.

Since the computation of the fully optimal strategy is prohibitively technical for practitioners, we propose a quasi-optimal strategy involving only a static constrained quadratic programme (CQP), easily implementable in a spreadsheet. We show numerically that the CQP strategy is practically indistinguishable from the optimal investment in terms of its welfare implications. We provide an explicit formula which helps visualize the lifestyling effect and further lowers the technical barrier towards its implementation. Furthermore, we compare the CQP strategy with other frequently used heuristic strategies. In order to confirm the accuracy of the proposed strategy, we perform a robustness analysis within a reasonable set of parametric settings.

Finally, we apply the CQP strategy to the system of retirement savings in Slovakia. In this particular situation, we are only dealing with the possibility of splitting the contributions between two types of funds, one of which is equity (unsecured) and one bond (guaranteed) pension fund. The strategy is tested on a typical lifecycle for different wage profiles. The accumulated savings are recalculated (using a standard annuity formula) to the replacement rate in order to assess the adequacy of the savings level.
An exponential time differencing method for American option pricing problems

Vera N. Egorova  
Universidad de Cantabria, Av. de los Castros, s/n, 39005 Santander – Spain  
Email: vera.egorova@unican.es

Rafael Company  
Universitat Polit`ecnica de Valncia, Camino de Vera, s/n, 46022 Valencia – Spain  
Email: rcompany@imm.upv.es

Lucas J´odar  
Universitat Polit`ecnica de Valncia, Camino de Vera, s/n, 46022 Valencia – Spain  
Email: ljodar@mat.upv.es

Abstract. The exponential time differencing (ETD) method is a promising approach for solving time-dependent partial differential equations (PDEs) in many fields. In computational finance, the price of an American option is described as the solution of a free boundary PDE problem that is to be solved then numerically. In present paper, semi-discretisation is used to transform the nonlinear PDE to a system of ordinary differential equations (ODEs). A numerical solution is constructed by using ETD and matrix quadrature rules.

1. Introduction. The most widely used in the field of computational finance Black-Scholes model provides a simple computable pricing formula of European option in an idealistic market. For American options, the closed form solution does not exist due to the early exercise opportunity. It makes the problem of American option pricing be a challenging task and a question of great interest in the field.

Due to the early exercise privilege, American put option pricing problem is described by the following free boundary PDE:

\[
\frac{\partial P}{\partial \tau} = \sigma^2 S^2 \frac{\partial^2 P}{\partial S^2} + r S \frac{\partial P}{\partial S} - r P, \quad S > B(\tau), \quad 0 < \tau \leq T,
\]

together with the initial and boundary conditions:

\[
B(0) = E, \quad P(S,0) = \max(E - S,0), \quad S \geq 0, \quad \frac{\partial P}{\partial S}(B(\tau),\tau) = -1,
\]

\[
P(B(\tau),\tau) = E - B(\tau), \quad \lim_{S \to \infty} P(S,\tau) = 0, \quad P(S,\tau) = E - S, \quad 0 \leq S < B(\tau),
\]

where \( \tau = T - t \) denotes the time to maturity \( T \), \( S \) is the price of the asset, \( P(S,\tau) \) is the option price at time \( \tau \), \( B(\tau) \) is the unknown early exercise boundary, \( \sigma \) is a volatility of the asset, \( r \) is the risk free interest rate and \( E \) is the strike price.

Previous research (see [2] and references therein) has suggested to apply the front-fixing transformation to treat the free boundary

\[ x = \ln \frac{S}{B(\tau)}, \quad p(x,\tau) = \frac{P(S,\tau)}{E}, \quad S_t(\tau) = \frac{B(\tau)}{E}, \]

that leads to the nonlinear PDE problem

\[
\frac{\partial p}{\partial \tau} = \frac{\sigma^2}{2} \frac{\partial^2 p}{\partial x^2} + \left( r - \frac{\sigma^2}{2} + \frac{S_t(\tau)}{S_t} \right) \frac{\partial p}{\partial x} - rp, \quad x > 0, \quad 0 < \tau \leq T,
\]

\[
S_t(0) = 1, \quad p(x,0) = 0, \quad x \geq 0, \quad \frac{\partial p}{\partial x}(0,\tau) = -S_t(\tau), \quad p(0,\tau) = 1 - S_t(\tau), \quad \lim_{x \to \infty} p(x,\tau) = 0,
\]

\]

[115]
Assuming that equation (5) holds at \( x = 0 \), it follows that

\[
\frac{\sigma^2}{2} \frac{\partial^2 p}{\partial x^2}(0, \tau) + \frac{\sigma^2}{2} S_t(\tau) - r = 0.
\]  

(7)

Problem (5)–(6) has to be solved numerically. In this work, we use semi-discretisation of the PDE to construct a system of nonlinear ODEs that is written in integral form. Then it is solved by the ETD method.

2. Numerical solution. Let us define a computational domain \( \Omega = [0, x_{\text{max}}] \) with uniformly distributed spatial nodes \( x_j \):

\[ x_j = jh, \ h = \frac{x_{\text{max}}}{M}, \ 0 \leq j \leq M. \]  

(8)

Then, denoting an approximate solution at the point \( x_j \) by \( p_j(\tau) \), the semi-discretisation of the equation (5) is obtained by applying the second order central difference approximation for the spatial derivatives, resulting in the system of ODEs on the time variable (for interior nodes, \( 1 \leq j \leq M - 1 \))

\[
p_j' = \frac{\sigma^2}{2} \frac{p_{j+1} - 2p_j + p_{j-1}}{h^2} + \left( r - \frac{\sigma^2}{2} + \frac{S'_t}{S_t} \right) \frac{p_{j+1} - p_{j-1}}{2h} - rp_j.
\]

(9)

The values at the boundaries of the computational domain are derived from (6) as follows

\[
p_0(\tau) = 1 - S_t(\tau), \quad p_M(\tau) = 0.
\]

(10)

From (6) and (7), by using the central finite difference of the second order one gets

\[
p_1(\tau) = \alpha - \beta S_t(\tau), \quad \alpha = 1 + \frac{rh^2}{\sigma^2}, \quad \beta = 1 + h + \frac{h^2}{2}.
\]

(11)

Deriving (11) with respect to \( \tau \), we obtain ODE:

\[
p_1'(\tau) = -\beta S'_t(\tau).
\]

(12)

From (10) and (11) the connection between \( p_0(\tau) \) and \( p_1(\tau) \) can be expressed as follows

\[
p_0(\tau) = \frac{1}{\beta} (\beta - \alpha + p_1(\tau)) \quad \tau > 0.
\]

(13)

The nonlinear system (9) contains \( M - 1 \) ODEs for \( M \) unknown functions \( (p_j(\tau), j = 1, \ldots, M - 1 \) and \( S_t(\tau) \)). To exclude \( S_t(\tau) \) from the system, the equations (12) and (11) are used to obtain the following relation

\[
\frac{S'_t}{S_t} = -\frac{p'_1}{\alpha - p_1}.
\]

(14)

Hence, taking into account (13) and (14), system (9) takes the vector form

\[
p' = A\bar{p} + \Phi(\bar{p}),
\]

(15)

where \( \bar{p} = [p_1, \ldots, p_{M-1}]^T \), \( A \) is a \((M - 1) \times (M - 1)\) tridiagonal matrix, \( \Phi(\bar{p}) \) is the vector that contains nonlinear part of the equations. Let us introduce constant coefficients

\[
a_{\pm 1} = \frac{\sigma^2}{2h^2} \pm \left( r - \frac{\sigma^2}{2} \right) \frac{1}{2h}, \quad a_0 = -\frac{\sigma^2}{h^2} - r,
\]

(16)

then matrix \( A \) takes the form

\[
A = \begin{pmatrix}
0 & 0 & 0 & \ldots & 0 & 0 & 0 \\
a_{-1} & a_0 & a_1 & \ldots & 0 & 0 & 0 \\
\vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots \\
0 & 0 & 0 & \ldots & 0 & a_{-1} & a_0
\end{pmatrix}.
\]

(17)
Denoting relation (14) by $-\phi$, vector $\Phi(\bar{p})$ of nonlinear part of the system (15) can be written as follows

$$\Phi(\bar{p}) = \begin{pmatrix}
\phi(\alpha - p_1) \\
-\phi_{p_2 - p_1} \\
\cdots \\
-\phi_{p_{j+1} - p_j} \\
\cdots \\
\phi_{\|p_{M+1} - p_M\|} \\
\end{pmatrix}$$

(18)

Finally, system (15) is solved numerically employing the ETD method [3]. For the temporal discretisation we fix the time step $k = \frac{T}{N}$, so that $t^n = nk$, $n = 0, \ldots, N$. Then the system of ODE (15) in some given interval $\tau \in [t^n, t^{n+1}]$ is equivalent to the vector integral equation (see Section 2.1 of [3]):

$$\bar{p}(t^{n+1}) = e^{Ak} \bar{p}(t^n) + \int_0^k e^{As} \Phi(\bar{p}(t^{n+1} - s))ds.$$  

(19)

We propose a first explicit approximation of the integral in (19) by replacing $\bar{p}(t^{n+1} - s)$ by the known value $\bar{p}(t^n)$ that corresponds to $s = k$. The local truncation error of such approximation is $O(k^2)$ [3]. Then (19) takes the form

$$\bar{p}(t^{n+1}) = e^{Ak} \bar{p}(t^n) + \left( \int_0^k e^{As} \Phi(\bar{p}(t^n)) \right).$$  

(20)

3. Results. In this Section, we present the numerical results of the proposed method performed by using MatLAB R2018a software.

The American option pricing problem considered in [4] takes the following parameters

$$r = 0.1; \quad \sigma = 0.2; \quad E = 100; \quad T = 1.$$  

(21)

For the transformed problem (5)–(6) with parameters (21) we set $x_{\text{max}} = \ln 3$, that corresponds to the choice $S_{\text{max}} = 3E$ at the initial moment. The proposed method is compared with the finite-difference method of [2] and with results of [4]. The numerical solution is constructed with $h = 10^{-2}$ and $k = 10^{-4}$.

The numerical solution of the problem is constructed as described above by using the ETD method. Then the inverse transform is applied to the solution in order to obtain the option price at the moment $\tau = T$, i.e. at the moment of signing the contract. Optimal stopping boundary $S_f(\tau)$ is given in 2 comparing with the results of the proposed in [2] FDM with logarithmic front-fixing transformation.

The results of this study indicate that it is a promising approach that can be extended to other scenarios as American call options and models taking into account jumps in the underlying asset becoming in partial integro-differential equations.

References


An exponential time differencing method for American option pricing problems

Figure 1: Prices of American option with parameters (21) as initial conditions ($\tau = 0$) and numerical solution by the proposed ETD method with $h = 0.01$, $k = h^2$.

Figure 2: Optimal stopping boundary by the proposed method (ETD) with $h = 0.01$, $k = h^2$ and the FDM from [2] for the problem with parameters (21).


Minisymposium 9:
Young researchers minisymposium on computational finance

Organized by
José G. López-Salas
Quasi-Regression Monte-Carlo scheme for semi-linear PDEs and BSDEs

EMMANUEL GOBET
ÉCOLE POLYTECHNIQUE, PALAISEAU, FRANCE

JOSÉ G. LÓPEZ-SALAS
UNIVERSIDADE DA CORUÑA, DEPARTMENT OF MATHEMATICS, 15071 A CORUÑA, SPAIN
EMAIL: jose.lsalas@udc.es

CARLOS VÁZQUEZ
UNIVERSIDADE DA CORUÑA, DEPARTMENT OF MATHEMATICS, 15071 A CORUÑA, SPAIN
EMAIL: carlosv@udc.es

Abstract. We present a novel quasi-regression Monte Carlo algorithm in order to approximate the solution of discrete time backward stochastic differential equations, and we analyze the convergence of the proposed method. The algorithm also approximates the solution to the related semi-linear parabolic partial differential equation obtained through the well known Feynman-Kac representation. For the sake of enriching the algorithm with high order convergence a weighted approximation of the solution is computed and appropriate conditions on the parameters of the method are inferred. With the challenge of tackling problems in high dimensions we propose suitable projections of the solution and efficient parallelizations of the algorithm taking advantage of graphics processing units (GPUs).
A ranking methodology for market making activity

Outhmane Mounjid
École Polytechnique, Route de Saclay, 91128 Palaiseau, France
Email: outhmane.mounjid@polytechnique.edu

Mathieu Rosenbaum

Pamela Saliba

Abstract. The goal of this talk is to present a methodology enabling us to assess the contribution of the various market participants to market quality. To do so, we use a generalized order book model involving both state dependence and memory effects for the dynamics of the order flows. In this framework, we are able to compute the effect of each participant on important market quantities such as volatility or liquidity. This leads us to the possibility to rank market makers in term of the effect of their activity on market quality.
Some neural network-based schemes to solve high-dimensional nonlinear PDEs

CÔME HURÉ
UNIVERSITÉ PARIS DIDEROT, PARIS, FRANCE
Email: hure@lpsm.paris
HUỲ%H PHAM
XAVIER WARIN

Abstract. We propose new neural network-based schemes to solve high dimensional nonlinear partial differential equations (PDEs). Our algorithm rely on the classical backward stochastic differential equation (BSDE) representation of PDEs, to estimate simultaneously the solution and its gradient by neural networks. These approximations are performed at each time step from the minimization of loss functions defined recursively by backward induction. The methodology is extended to variational inequalities arising in optimal stopping problems. We analyze the convergence of the deep learning schemes and provide error estimates in terms of the universal approximation of neural networks. Numerical results show that our algorithms give very good results till dimension 50 (and certainly above), for both PDEs and variational inequalities problems. For the PDEs resolution, our results are very similar to those obtained by the recent method in [1] when the latter converges to the right solution or does not diverge. Our numerical tests indicate that the proposed methods are not stuck in poor local minima as it can be the case with the algorithm designed in [1], and no divergence is experienced.

References

Optimal relative consumption by competing with past spending maximum

Shuoqing Deng  
Université Paris Dauphine, France  
Email: deng@ceremade.dauphine.fr

Xun Li  
Huyen Pham  
Xiang Yu

Abstract. We consider the financial problem of optimal investment and consumption where the investors utility function depends on not only the consumption but rather on the consumption level compared with the past spending maximum. In terms of financial interpretation, this means a type of habit formation where the investors satisfactory level depend on its past living standard. We provide the closed form expressions for the optimal consumption and investment strategy in the dual formulation and prove that the candidate value function is indeed verified.
Finite variance unbiased estimation of stochastic differential equations

Ankush Agarwal
Email: ankush.agarwal@glasgow.ac.uk

Emmanuel Gobet

Abstract. We develop a new unbiased estimation method for Lipschitz continuous functions of multi-dimensional stochastic differential equations with Lipschitz continuous coefficients. This method provides a finite variance estimator based on a probabilistic representation which is similar to the recent representations obtained through the parametrix method and recursive application of the automatic differentiation formula. Our approach relies on appropriate change of variables to carefully handle the singular integrands appearing in the iterated integrals of the probabilistic representation. It results in a scheme with randomized intermediate times where the number of intermediate times has a Pareto distribution.
Conditional Monte Carlo learning for diffusions

LOKMAN ABBAS-TURKI
UNIVERSITÉ PARIS SORBONNE, FRANCE
EMAIL: lokmane.abbas_turki@upmc.fr

Abstract. We present a new algorithm based on a One-layered Nested Monte Carlo (1NMC) to simulate functionals $U$ of a Markov process $X$. The main originality of the proposed method comes from the fact that it provides a recipe to simulate $U_{t \geq s}$ conditionally on $X_s$. This recipe can be used for a large number of situations including; Backward Stochastic Differential Equations (BSDEs), Reflected BSDEs, risk measures and beyond. In contrast to previous works, our contribution is based on a judicious combination between regression and 1NMC used for localization purpose. Although the procedure is initially developed for diffusions, we also explain in which situations it can be applied to jump diffusions. The generality, the stability and the iterative nature of this algorithm, even in high dimension, make its strength. It is of course heavier than a straight Monte Carlo (MC), however it is far more accurate to simulate quantities that are almost impossible to simulate with MC. Moreover, the parallel suitability of 1NMC makes it feasible in a reasonable computing time. In this presentation, we explain this algorithm, study its convergence and complexity. We also provide various numerical examples with a dimension bigger than 100 that are executed in few minutes on one Graphics Processing Unit.
Numerical approximation of BSDEs with polynomial growth driver

ARNAUD LIONNET
UNIVERSITY OF BIRMINGHAM, UNITED KINGDOM
EMAIL: a.lionnet@bham.ac.uk

Abstract. Backward Stochastic Differential Equations (BSDEs) provide a systematic way to obtain Feynman–Kac formulas for linear as well as nonlinear partial differential equations (PDEs) of parabolic and elliptic type, and the numerical approximation of their solutions thus provide Monte-Carlo methods for PDEs. BSDEs are also used to describe the solution of path-dependent stochastic control problems, and they further arise in many areas of mathematical finance.

In this talk, I will discuss the numerical approximation of BSDEs when the nonlinear driver is not Lipschitz, but instead has polynomial growth and satisfies a monotonicity condition. The time-discretization is a crucial step, as it determines whether the full numerical scheme is stable or not. Unlike for Lipschitz driver, while the implicit Bouchard–Touzi–Zhang scheme is stable, the explicit one is not and explodes in general. I will then present a number of remedies that allow to recover a stable scheme, while benefiting from the reduced computational cost of an explicit scheme. I will also discuss the issue of numerical stability and the qualitative correctness which is enjoyed by both the implicit scheme and the modified explicit schemes. Finally, I will discuss the approximation of the expectations involved in the full numerical scheme, and their analysis when using a quasi-Monte Carlo method.
Abstract. The goal of this work is to parallelise the multistep method for the numerical approximation of the Backward Stochastic Differential Equations (BSDEs) in order to achieve a high accuracy and a reduction of the computation time as well. In the multistep scheme, the computations at each grid point are independent and this fact motivates us to select GPU computing using CUDA. In our investigations we identify performance bottlenecks and apply appropriate optimization techniques for reducing the computation time, using a uniform domain. Finally, various numerical examples including financial applications are provided to demonstrate the achieved acceleration on GPU for solving BSDEs.
Contributed talks
Incorporating taxation in the valuation of variable annuity contracts: the case of the guaranteed minimum accumulation benefit

JENNIFER ALONSO-GARCÍA
DEPARTMENT OF ECONOMICS, ECONOMETRICS AND FINANCE, UNIV. OF GRONINGEN, THE NETHERLANDS & CEPAR
Email: j.alonso-garcia@rug.nl

MICHAEL SHERRIS
SCHOOL OF RISK AND ACTUARIAL STUDIES, UNSW SYDNEY, AUSTRALIA & CEPAR
Email: m.sherris@unsw.edu.au

SAMUEL THIRURAJAH
SCHOOL OF RISK AND ACTUARIAL STUDIES, UNSW SYDNEY, AUSTRALIA & CEPAR
Email: samuel.thirurajah@student.unsw.edu.au

JONATHAN ZIVEYI
SCHOOL OF RISK AND ACTUARIAL STUDIES, UNSW SYDNEY, AUSTRALIA & CEPAR
Email: j.ziveyi@unsw.edu.au

Abstract. This paper considers variable annuity contracts embedded with guaranteed minimum accumulation benefit riders (GMABs) promising the return of the premium paid by the policyholder, or a higher stepped up value at the end of the investment period in the absence of taxation. The valuation framework is presented in partial differential form which is solved for fair fees with the aid of the method of lines algorithm. Two taxation regimes are considered; one where capital gains can offset losses and a second where gains cannot offset losses, reflecting the institutional arrangements in Australia and the U.S., respectively. When capital losses cannot offset gains, the separation between the insurer and policyholder contract valuation curves increases as tax increases, reflective of the increasing value of the contract to the governments. If capital losses can be offset by capital gains, policyholder’s optimal surrender behavior changes to the benefit of the insurer, reducing the insurer liabilities for any fee rate they choose to charge.

1. Introduction. Variable annuities (VAs) are notoriously popular in the US where the net asset value was approximately $1.95 trillion as of March 2018 [2]. On the other hand, there is a very thin market for VAs in Australia, Europe and the rest of the world. The variable annuity market is relatively immature in Australia. In Europe, the VAs’ market was worth 188 billion in 2010 [3]. However, after the Global Financial Crisis, their popularity decreased and various life insurers stopped offering such contracts[7]. In Japan, the VA market grew from a market of less than $1 billion in 2000 to over $50 billion, due to financial deregulation in the late 90s [4]. VAs are among the few assets which grow tax-deferred within the US and Australia.

The greater part of existing literature has focused on risk-neutral valuation of VA contracts using a variety of techniques without considering income and wealth tax. For example, Bauer et. al. (2008) provide a universal pricing framework for various riders embedded in variable annuities using Monte Carlo based algorithms under the geometric brownian motion (GBM) framework[1].

2. Methodology. The GMAB rider discussed in this paper involves a policyholder entering into a VA contract by investing an initial amount $x_0$ into a mutual fund. Upon maturity of the contract, the policyholder is promised the greater of the minimum guarantee, $G$, and the fund value. In order to finance this guarantee, the insurer charges a continuously compounded fee at rate $q$ which is deducted as a percentage of the fund. Assume that the underlying $(S_t)$ follows a standard geometric Brownian
Incorporating taxation in the valuation of VA contracts: the case of the GMAB

motion (GBM) under the risk-neutral measure. The VA fund value \(x_v\) evolves according to the following dynamics:

\[ dx_v = (r - q)x_v dt + \sigma x_v dW_v. \]

(1)

Furthermore, although the income of the policyholder is taxable, they are not taxed until early surrender or maturity. This is consistent with the US tax system where such contracts are tax deferred and switching funds is tax free. The taxable income of the policyholder at maturity can be re-written:

\[ \max(G, x_T) - x_0 - C_0 - y(T)_+, \]

(2)

where \(C_0\) is any additional up front cost incurred on top of the premium \(x_0\), \([z]_+ = \max(z, 0)\) and \(y(T)\) is the numerical value of the total fees paid from the fund throughout the life of the contract. This reflects tax system in which the policyholder is only taxed upon withdrawal from the account reflecting its tax-deferred nature of VAs in the US and Australia. Moreover, we assume that the guarantee fees paid constitute the cost base of the product. Hence, we subtract out the fees paid \(y(T)\) for tax calculation purposes. In general, \(y(\nu) = \int_0^\nu q \cdot x(s)ds\) is the cumulative fees paid at time \(\nu\). We assume two tax treatments for losses. The first assumption is that taxable income is nonnegative, that is, capital losses incurred by the VA account cannot be offset by gains from other investments. This is reflective of the tax situation in the US [5]. The case when losses offset gains is considered as an extension that reflects the tax treatment in Australia.

Let \(u^p(x, y, \nu)\) be the value of the GMAB contract for the policyholder where \(x\) represents the fund value, the accumulative fees paid is \(y(\nu) = \int_0^\nu q \cdot x(s)ds\) and the time elapsed since the inception of the contract is \(\nu\). Using hedging arguments and Ito’s Lemma, it can be shown that the value of the contract is the solution to the following partial differential equation:

\[ \frac{1}{2} \sigma^2 x^2 u_{xx} + xq \cdot u_y + (r - q) \cdot xu_x - ru_y - u^p_t = 0. \]

(3)

Note that we applied the transformation \(t = T - \nu\) and \(t\) represents the time to maturity on the contract. The PDE(3) is solved subject to the initial and following boundary conditions:

\[ u^p(x, y, 0) = \max(x, G) - \tau \left[ \max(x, G) - y - x_0 - C_0 \right]_+ \]

(4)

\[ u^p(s(t, y), y, t) = s(t, y) \gamma_t - \tau [s(t, y) \gamma_t - y - x_0 - C_0]_+, \]

(5)

\[ u^p(0, y, t) = (G - \tau [G - y - x_0 - C_0]_+)e^{-rt}, \]

(6)

\[ u^p_y(s(t, y), y, t, \gamma_t) = \gamma_t - \tau \gamma_t I \{ s(t, y) \gamma_t - y - x_0 - C_0 > 0 \}, \]

(7)

where \(\gamma\) is the proportion that the policyholder is allowed to keep subsequent to surrender, \(x_0\) is the initial fund value (that is, the ‘premium’), \(C_0\) is the initial upfront cost, \(\tau\) is the tax rate, \(G\) is the guarantee amount and \(s(t, y)\) is the minimum fund value to trigger surrender, given that there are still \(t\) years to maturity and \(\gamma\) fees have been paid so far. \(s(t, y)\) is a free boundary that must be computed along with the valuation solution \(u\). The first two boundary conditions, Equations (4) and (5), represent the post-tax payoff at maturity or upon surrender, which occurs when the fund value \(x\) exceeds \(s(t, y)\). Equation (6) is the present value of the taxable income at maturity when the fund value is 0 as given by Equation (2). Hence, in this case the guarantee is paid with all certainty and the payoff is deterministic. Also note that it is usually the case that \(G < y + x_0 + C_0\), so there is no tax and the right hand side simplifies to \(Ge^{-rt}\). The final boundary condition, Equation (7), enforces the continuity of \(u^p\) at the free boundary \(x = s(t, y)\).

For the case in which capital losses on the GMAB rider can be used to offset other income sources, as is the case for nonqualified plans in the US and Australian variable annuities. Mathematically, this entails the following replacement

\[ \tau(\gamma x - y - x_0)_+ \rightarrow \tau(\gamma x - y - x_0), \]

\[ ^1\text{We reserve the use of t for the time to maturity and } \tau \text{ for the tax rate. Therefore, we use } \nu \text{ to denote the time elapsed since the inception of contract}. \]
in the boundary conditions (4)-(7). In this case, however, the additional condition at \( y = Y \) will be different as high levels of fees paid do not tend to the no-taxation case. In order to solve the PDE (3) subject to the initial and boundary conditions (4)-(7), we utilise the numerical method of lines algorithm [6].

3. Findings. An insightful way of visualising policyholder surrender behaviour is to plot the surrender boundaries. Figure 1 below illustrates the surrender boundaries under different tax regimes: in the absence of tax, when the tax paid is allowed to be negative (‘offset’) and when the tax paid must always be positive (‘no offset’).

![Figure 1: Impact of tax regime on surrender surface](image)

In Figure 1 we present the surrender boundaries under various tax systems. Due to the fair fee being low for the no offset case, the surrender boundary ends up being much higher than the other two cases, with the exception of the ‘valley of surrender’. In contrast, when offsets are allowed the surrender boundary is monotonically decreasing in \( y \), reflecting the incentive of the policyholder to surrender sooner if they have already paid a large amount of fees since they will be eligible for a large deduction on their taxes.

We adopt an exponential decreasing surrender fee structure such that \( \gamma_t = e^{-\kappa t} \). This parameter \( \kappa \) is also varied in the sensitivity analysis in Table 1 above.

Table 1 demonstrates that when the parameters \( r \) and \( \sigma \) are varied within realistic ranges, there is a significant impact on the fair fees from the policyholder and insurer perspectives. A higher \( r \) is accompanied by a lower fee rate. Since policyholders can obtain a greater return in the risk-free market, they may be less interested to enter the contract for the same level of maturity guarantee. A greater value for the volatility \( \sigma \) also implies an increase in the fair fee, since the guarantee has greater value in a more volatile market.

4. Conclusion. In this paper we have illustrated the impact of different tax systems on policyholder behaviour and the implications for insurers. We have also examined contract values to highlight the
Table 1: Sensitivity analysis: analysis of the impact of $r$, $\sigma$ and $\kappa$ on policyholder ($q^p,^*$) and insurer ($q^i,^*$) fair fees.

<table>
<thead>
<tr>
<th>Losses offset gains</th>
<th>Capital gains only</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\sigma=0.20$, $r=0.03$</td>
<td>$\kappa=0.005$</td>
</tr>
<tr>
<td>$r=0.025$</td>
<td>$q^p,^*$</td>
</tr>
<tr>
<td>$\tau=0$</td>
<td>3.35</td>
</tr>
<tr>
<td>$\tau=0.225$</td>
<td>1.97</td>
</tr>
</tbody>
</table>

impact of various fee rates being charged on the tax revenue that is collected by the government. We presented a method of lines approach that enables us to efficiently determine optimal surrender boundaries, contract values and fair fee rates. A model incorporating taxation is able to more accurately capture policyholder surrender behaviour.

We have formulated the valuation problem of a GMAB contract from the policyholder perspective as a free boundary problem which can be solved using the method of lines. We have used the method of lines to solve the free boundary problem to value the contract from the insurer’s perspective. The corresponding policyholder fair fee and insurer fair fee has also been computed. The numerical results illustrate that $r$, $\sigma$ and $\kappa$ have an impact on optimal surrender. This impact is contingent on how much fees the policyholder has cumulatively paid so far and also on the particular taxation system, such as whether or not capital losses are permitted.

References


HPC framework for multiple evaluation of supervised learning models and a famous expert for stock screening using fundamental factors

POL ALVAREZ
BARCELONA SUPERCOMPUTING CENTER, BARCELONA, SPAIN
Email: pol.alvarez@bsc.es

ARGIMIRO ARRATIA
DEPT. OF COMPUTER SCIENCE, BARCELONA TECH (UPC), SPAIN
Email: argimiro@cs.upc.edu

Abstract. We present HPC.FASSR, a High Performance Computation Fundamental Analysis Stock Screener and Ranker system, build on PyCOMPSs to compare the performance of various supervised learning algorithms like neural networks, random forests, support vectors machines, or AdaBoost, and well-known human expert trader’s criteria for selecting stocks based on fundamental factors. The parallelization of HPC.FASSR with PyCOMPSs allow us to explore a huge number of configurations in a short time. Our experiments are extensive. Here we only report the performance of the machines against the human expert. The best models do exceed the expert; however the expert is able to beat the average model for each class of learning machine. All tests have been executed in MareNostrum 4, the main supercomputer in the Barcelona Supercomputing Center.

1. Data and Methods.

1.1. Overview. The HPC.FASSR system involves the integration of many techniques. Figure 1 shows the pipeline of processing and transforming the data from source to results. This is divided into four major stages. The first stage is responsible for downloading the desired data from the API and merging the financial statements. Next follows the data processing stage where the final datasets are created by joining fundamental indicators, generating the features and targets, computation of z-scores, oversampling to weekly periods, and filtering missing values. At the third stage all models considered are trained with all of their possible parametrizations, and every available dataset. At the fourth stage each trained model is tested simulating a real trading environment.

1.2. Variables and targets. For this study we collected quarterly financial statements and price data from the 505 companies listed in the S&P500 American index. Data were downloaded from the Intrinio platform1. This selection does not include old S&P500 constituents so the results may suffer from survivorship bias. However, as all models are tested with the same dataset, the comparisons among them are valid. From the financial statements we compute the following fundamental indicators to be used by the learning models and the human expert: Earnings per share; Price to Earnings; Price to Book; Price to Revenue; Dividends to price; Dividend payout ratio; Return on Equity; Return on Invested Capital; Return on Assets; Asset turnover; Inventory turnover; Profit margin; Debt ratio; Earnings to Interest Expense; Current Ratio; Revenue; Working Capital; Earnings growth; and Book value per share. If the financial statements of a company in a given quarter lacks any field required to compute a fundamental indicator, the company is removed from that quarter data. This results in an average of 190 companies per period. Since financial statements are available on a quarterly basis, one ends up with a small amount of yearly training samples. To mitigate this problem we applied oversampling as follows. Most

---

1. This selection does not include old S&P500 constituents so the results may suffer from survivorship bias. However, as all models are tested with the same dataset, the comparisons among them are valid.
2. From the financial statements we compute the following fundamental indicators to be used by the learning models and the human expert: Earnings per share; Price to Earnings; Price to Book; Price to Revenue; Dividends to price; Dividend payout ratio; Return on Equity; Return on Invested Capital; Return on Assets; Asset turnover; Inventory turnover; Profit margin; Debt ratio; Earnings to Interest Expense; Current Ratio; Revenue; Working Capital; Earnings growth; and Book value per share. If the financial statements of a company in a given quarter lacks any field required to compute a fundamental indicator, the company is removed from that quarter data. This results in an average of 190 companies per period. Since financial statements are available on a quarterly basis, one ends up with a small amount of yearly training samples. To mitigate this problem we applied oversampling as follows. Most
of the indicators relate fundamental data with stock price, the latter being available on a daily period. Hence, to generate new samples for a given date, we compute the indicator using the previous quarter fundamental data adjusted to the stock price of the given date. For the indicators which are not related to stock price (e.g., revenue) we use the latest available value.

Targets: For regression we used the simple return, i.e. \( y_{i,s} = \frac{r_{i,s}}{r_{i-1,s}} \). For classification we categorized the returns into long, neutral, short. The thresholds used were ±3%. Stocks with a return higher than 1.03 are labelled as long, stocks with a return under 0.97 are labelled as short, the rest as neutral.

1.3. Forecasting models. We used scikit-learn\(^2\) implementations of the most popular machine learning predictors in literature: Neural Networks, Support Vector Machines, Random Forests, and AdaBoost. The system is designed to work with models that follow scikit-learn model’s API so users can integrate their custom models by just implementing the common interface. Moreover, the system supports models that can run in more than one CPU. Using the standard scikit-learn \( n\_jobs \) parameter and PyCOMPSs \@constraint decorator, a model can be trained with any desired number of CPUs. This highly increases the performance of embarrassingly parallel models like random forests.

The human expertise criteria that we programmed as a rule-based decision algorithm, and that is to compete against the above classifiers, is that of famous value investor Benjamin Graham. His rules for selecting value stocks, updated for inflation to present time, are the following (see [1, §6.2] for details):

- the company should have at least USD 1.5B in revenues;
- a 2-to-1 assets to liability ratio;
- positive earnings in each of the past ten years;
- uninterrupted payment of dividends;
- a 3% of annual average growth in earnings;
- moderate P/E and P/B ratios.

1.4. Experimental set-up and evaluation. We have evaluated the performance of all models taking long-short positions, trading every month, quarter, semester or year, through different regimes, and with different trading strategies. Additionally we considered training with information normalised as z-scores grouped by industry, and with either one or two years of training data. We also evaluate the parallelization performance by running the experiments with different amounts of computing resources.

To evaluate the performance of the models, the trading module simulates a paper trading environment to do backtesting. Essentially, the trading time interval is divided into trading sessions. Figure 2 shows the execution flow of a single trading session. For each trading session, we apply the model to the input data to obtain a set of next period return predictions. For regressors each prediction is converted through a selection function to the class of investment position for the stock according to a fixed threshold value

\(^2\)https://scikit-learn.org/stable/
Figure 2: Example of a single iteration of the trading evaluation. This process is repeated for each trading session and the final portfolio and available money are used as input for the next.

for the expected return: long position when the predicted return is over 3%, short when predicted return is less than -3%, and neutral otherwise. For classifiers the positions are given by the prediction. The last step it to create a new portfolio using a trading strategy. The trading strategy takes as input the long-short recommendations, the previous portfolio, and the available money to decide which old positions should be closed and which new ones should be bought. After evaluating all models, we choose the best one according to a desired target metric. Some target metric examples are: the total profits, the average invested money, or a combination of them with the execution time (if we want a trade-off between accuracy and performance). All the performance tests have been executed in the MareNostrum 4 supercomputer. MareNostrum 4 has 3456 computing nodes housing a total of 165,888 cores, 390 Terabytes of main memory, and a peak performance of eleven thousand trillion operations per second (11.5 Petaflops).

1.5. PyCOMPSs. PyCOMPSs [2] is a framework aimed to ease the development of parallel and distributed Python applications. The main purpose is to abstract users from both the infrastructure management and the data handling. PyCOMPSs is composed of two main parts: the programming model and the runtime.

1.5.1. Programming model. PyCOMPSs offers a sequential programming model paradigm where users only need to indicate which parts of the code can be executed in a distributed fashion.

To indicate which parts of the code execution can be distributed, PyCOMPSs takes advantage of the python function decorators. The basic PyCOMPSs decorator is @task() and it states that a given function should be treated as a task, the basic distributed execution block. In HPC.FASSR, each training and evaluation of a model is run as task. Tasks are run distributed and their parameters need to be instrumented to transfer both inputs and outputs between the nodes. All the inputs and outputs of tasks are placeholders to allow them to run asynchronously while the main code continues to be executed. PyCOMPSs offers many advanced functionalities, one of them, the @constraint() decorator, allows to specify hardware constraints at task-level (e.g., required CPUs to run a task).

1.5.2. Runtime. The PyCOMPSs framework is mostly a set of bindings to interact with the java runtime. The runtime has two main responsibilities: tasks’ dependency analysis and scheduling, and resources and data management. To execute the functions decorated as tasks in distributed the runtime needs to find the data dependencies between them. That information is then used to build a task dependencies graph which determines the order of the execution. Once tasks are free of dependencies in the graph, they are scheduled to be executed in some of the available computing resources.

With respect to the resources management, the runtime follows a master-worker paradigm. All available computing resources are represented uniformly as workers. The master node is responsible of executing the main code. When a task is called, the master schedules its execution (following the dependencies graph) to a given worker, transfers its required input, and, once it is completed, gathers its outputs, if any. Thanks to the pluggable connectors PyCOMPSs code is infrastructure-agnostic and
Multiple evaluation of learning models for stock screening

Figure 3: Comparison of Graham criteria and the best learning models. The values are the total USD at each time step (invested + free cash) per model.

Figure 4: Boxplot of the total returns of different parametrizations grouped by model.

can be run in a wide number of platforms without requiring any change (e.g. clouds, cluster and grid nodes, or docker/singularity containers).

2. Experimental Results. Here we only have space for reporting our experiments comparing the Graham criteria performance against all ML models. Graham’s criteria only applies long positions; so, in order to make the comparison fair, the ML models considered are restricted to trade only long positions. For this report we have trained around fifty thousand models, Figure 3 shows the returns evolution of the best combinations of parameters/dataset per machine model and Graham. Figure 4 shows box plots of the average return of all tested trading periods and parametrizations for each machine model and the Graham model. On the HPC side, we evaluated the scalability of HPC.FASSR running it with 1 node (48 CPUs) to 16 nodes (768 CPUs) and training an evaluating a test set of 1743 models, achieving a final speedup of x1.8, and training up to 800 models per minute.

As a general conclusion we note that not any model can beat Graham criteria, and in order for the automaton to beat the consistent gains obtained by following this expert advice, one has to make high investments in computational resources.

References


Abstract. We analyse in this work the performance of four popular implementations of stop-loss rules (fixed percentage barrier, average true range, relative strength index, triple moving average crossover) applied to asset prices whose returns are modelled with consideration of overnight gaps and acute momentary price drops (flash crashes). For this analysis we consider different models of asset returns: random walk, autoregressive and regime-switching models. As a general conclusion we find that, even when including overnight gaps and flash crashes in our price models, in rising markets stop-loss rules improve the expected risk-adjusted return according to most metrics, while improving absolute expected return in falling markets. Furthermore, we find that the fixed percentage stop-loss rule may be, in risk-adjusted terms, the most powerful among the rules that this work considers.

1. Introduction. A stop-loss order is an order an investor may place so that her position is liquidated the moment a certain pre-specified condition (set according to a stop-loss rule) is met by market dynamics. The purpose of setting a stop-loss is to cut losses and consolidate gains, being this a most basic yet popular tool for risk management.

Several researchers have looked at the problem of discerning the effectiveness of stop-loss orders in various contexts, see for instance [1], [3], [4], [5]. None of these works, and others that we surveyed, considers the oftentimes observed large variation in the price across non trading hours. This clearly affects the correct triggering of a stop-loss rule due to the by-passing of the established stopping-time. The present work contributes to the analysis of the value of stop-loss rules in various novel aspects. First, given that an accurate assessment of the value of stop-loss rules requires full consideration of the real behavior of asset returns, we take into consideration two important anomalies: 1) overnight gaps in prices (jumps from the Closing price of one day to the Open price of next trading day) as an additional feature that we include in the implementation of various well-accepted models for the behavior of asset prices; 2) flash crashes, an extreme market event that consists in a sharp price drop caused by a variety of non-fundamental reasons that is offset in the next minutes or hours.

2. Stop-loss rule and stop-loss policy. Given an asset $A$, we consider a stop-loss rule as a decision criterion $SL_t$ on the price history of $A$ such that at every time $t$ it takes the value 0 if the criterion is not satisfied or 1 otherwise. We denote by $\nu(SL_t)$ the value (0 or 1) of $SL_t$. Then, given a time horizon $T$ and an initial time instant $t_0 < T$ at which we buy a risky asset $A$, a stop-loss policy is a risk management measure that depending on the value of a stop-loss rule it signals to exit the investment at a certain time $t$, $t_0 \leq t \leq T$, where then we shall sell all of our holdings in $A$ and use the proceeds to go long on a risk-free asset $F$ up until time $T$.
In our experiments the risk-free asset $F$ is a AAA government bond, e.g., a 1-year U.S. government bond. As in [4], in our models we assume that the returns $\{r_t\}$ of the risky asset $A$ satisfy: The expected return $\mu$ of $A$ is greater than the risk-free rate $r_f$, and let $\pi := \mu - r_f > 0$ denote the risk premium of $A$. This property simply excludes the perverse case where the stop-loss policy adds value just because the risk-free asset that the investor transfers the capital to has a higher expected return than $A$. In our models we ensure this property holds by setting: i) the risk-free rate to the historical mean return of a 1-year U.S. government bond, which, during the period 1990-2017 has been 3.17%; ii) the expected return of our hypothetical asset is set to the average return of the NYSE during the same period (1990-2017), which is 6.14%. Assuming compounding, this implies an hourly return of 3.942924 $\times 10^{-5}$.

In our study we implement and analyze the performance of the following four stop-loss rules ($SL_t$):

**Fixed percentage barrier (%)**: $SL_t := P_t < P_{t-1}^{max}(1 - a)$, where $P_{t-1}^{max}$ is the highest price achieved until time $t-1$ and the positive quantity $a \cdot 100\%$ is the maximum percentage of that price $P_{t-1}^{max}$ that the investor is willing to lose in its position. Note that this corresponds to a mobile support barrier common among practitioners. The value of $a$ can be adjusted from historical data or be given a constant value based on a certain risk profile.

**Average True Range (ATR)**: $SL_t := P_t < P_{t-1} - \alpha ATR_d$, where $\alpha \in [1.5, 3], d \leq t$ and $ATR_d$ is a crude estimation of daily historical volatility in the recent past $ATR_d := \frac{1}{N} \sum_{i=0}^{N} T\bar{R}_{d-i}$, with $T\bar{R}_d := \max\{High_d - Low_d, High_d - Close_{d-1}, Close_{d-1} - Low_d\}$, $High_d$, $Low_d$ and $Close_d$ are the maximum, minimum and closing price at time $d$, and a typical value for $N$ is 14 days. Note that whilst $ATR_d$ is measured on a daily basis, $SL_t$ can be hourly or follow other higher time frequency.

**Relative Strength Index (RSI)**: The RSI is a momentum indicator that compares the magnitude of an asset’s recent gains and losses over a specified time period. It is defined as $RSI_t(w) := 100 \times \frac{100}{1 + RS(w)_t}$, where $RS(w)_t$ is the average gain of up periods divided by the average loss of down periods, during the specified time window of $w$ (usually in between 7 to 14 trading sessions). In this case the stop-loss rule is: $SL_t := RSI_t(w) \geq 70$.

**Triple Moving Average crossover (MA)**: A MA of $p$ periods is defined as $MA(p) := \frac{1}{p} \sum_{i=0}^{p-1} C_{t-i}$, where $C_t$ is the asset’s closing price at session $t$. The stop-loss rule in this case reads, for $a = 5$, $b = 20$ and $c = 70$, $SL_t := MA(a) < MA(b) \wedge MA(b) < MA(c)$. That is, the stop-loss rule at any time $t$ is given by two events that must take place simultaneously: the 5-periods MA is below the 20-periods MA, and the latter is in turn below the 70-periods MA.

2.1. **Performance metrics.** We will assess the impact of a stop-loss policy on investment performance with five metrics (the last two are new contributions): Expected return; Sharpe Ratio ($SR$); Sortino Ratio ($SOR$); Return-VaR ratio ($RVaR$, which is defined as the median return, in excess of the risk-free rate, divided by the median return minus the 1-year 5% VaR); Return-ES ratio ($RES$, defined as the median return, in excess of the risk-free rate, divided by the median return minus the 1-year 5% Expected Shortfall).

3. **Models of asset prices with overnight gap and flash crash.** The parameter estimation of our models are based on the following set of 30 NYSE stocks among the most liquid by 2017: (BAC, GE, PFE, S, F, C, T, JPM, WFC, HPQ, KO, AMD, MRK, XOM, JCP, ABX, RAD, GLW, VZ, JNJ, AIG, PG, DIS, HAL, XRX, KEY, BMY, SCHW, ABT, MO). We simulate hourly quotes (high-frequency modelling) running through several days, so can be affected by overnight gaps.

An overnight gap is defined as the difference between the Open price at day $t$, and the Close price at day $t-1$. Using our empirical data on 204,120 gaps (27 years worth of data on the 30 NYSE stocks), we find that the average gap is about 1.000463 (i.e. +0.0463%), and on average, gaps of any size above
a minimum threshold of 0.1% happen around 77.93% of the days. This is the frequency we used in our simulations. We modeled overnight gaps as a Weibull random variable. By considering overnight gaps we obtain a more reliable assessment of stop-loss performance.

A flash crash is a rapid, deep, and volatile fall in the price of a security, followed by a rebound, occurring within an extremely short period of time. A flash crash is modeled as a Bernoulli random variable with probability 0.05% per trading hour, which yields an expected number of such crashes of about one per year, and the magnitude of the flash crash (conditional on its occurrence), we model as a uniform random variable in the interval 5% - 35%. These values are in line with the flash crash episodes observed in the last decade.

These two important anomalies which clearly affect the performance of stop loss rules are incorporated to our price models. We consider the three most accepted models for risky asset’s returns in the literature about analysis of stop-loss rules [4, 5]. These are: random walk (with returns following a Generalized Normal distribution (GED)), ARMA model, and regime-switching (RS) model. We use a GARCH(1, 1) to model the volatility of returns. For the simulations, we consider hourly quotes from March 13 to March 24 2017 (both included), for each of the 30 stocks, which equates to 2070 hourly returns.

4. Experimental Results. We run Monte Carlo simulations, based on 5,000 repetitions for each stop-loss rule. Because we repeat each set of simulations 50 times, we are able to use the Welch’s t-test to assess the statistical significance of the difference in mean values that the stop-loss rules provide compared to Buy-and-Hold. We use the following notation to denote statistical significance: ns: p-value \( \geq 0.05 \), *: \( 0.01 \leq \text{p-value} < 0.05 \), **: \( 0.001 \leq \text{p-value} < 0.01 \), ***: \( \text{p-value} < 0.001 \).

Tables 1, 2, 3 and 4 show the out-of-sample results for the fixed percentage stop-loss rule, comparing stop-loss performance both without, with overnight gaps and further with flash crash, and for the MA crossover rule considering models with gaps and flash crashes (the results for the other two rules, RSI and ATR, are left out of this report).

As a general conclusion, both in the models without and with gap, we observe that all stop-loss rules provide a higher risk-adjusted return than Buy-and-Hold (B&H). The results are particularly remarkable for the fixed percentage rule with returns following a Regime-Switching (RS) model with gap, but no flash crash (Table 2): Sharpe Ratio is approximately multiplied by a factor of three, Sortino is multiplied by eight, while RVaR and RES are roughly multiplied by seven. The MA crossover-based stop-loss rule (Table 4) shows to be robust to flash crashes, which is an intuitive result, as a momentary sharp (yet realistic) price drop is unlikely to trigger the crossing of the short, medium and long-term moving averages. The RSI rule also showed to be robust to flash crashes (not shown here). Further we note that the usefulness of each stop-loss rule is apparent in all the considered models for returns. For further details see the extended version of our work [2].

References


## Models with overnight gap and flash crash to test stop-loss rules

### Table 1: Fixed percentage SL, without gap: Simulation results

<table>
<thead>
<tr>
<th>Strategy-Model</th>
<th>$\mathbb{E}[r]$</th>
<th>Sharpe</th>
<th>Sortino</th>
<th>RVaR</th>
<th>RES</th>
</tr>
</thead>
<tbody>
<tr>
<td>SL - GED</td>
<td>4.57%</td>
<td>0.81</td>
<td>2.02</td>
<td>0.85</td>
<td>0.75</td>
</tr>
<tr>
<td>B&amp;H - GED</td>
<td>7.09%</td>
<td>0.32</td>
<td>0.16</td>
<td>0.16</td>
<td>0.14</td>
</tr>
<tr>
<td>Difference</td>
<td>-2.52% ns</td>
<td>+0.49***</td>
<td>+1.86***</td>
<td>+0.69***</td>
<td>+0.61***</td>
</tr>
<tr>
<td>SL - ARMA</td>
<td>4.51%</td>
<td>0.82</td>
<td>2.09</td>
<td>0.86</td>
<td>0.76</td>
</tr>
<tr>
<td>B&amp;H - ARMA</td>
<td>6.82%</td>
<td>0.34</td>
<td>0.17</td>
<td>0.18</td>
<td>0.15</td>
</tr>
<tr>
<td>Difference</td>
<td>-2.31% ns</td>
<td>+0.48***</td>
<td>+1.92***</td>
<td>+0.68***</td>
<td>+0.61***</td>
</tr>
<tr>
<td>SL - RS</td>
<td>3.66%</td>
<td>0.82</td>
<td>3.20</td>
<td>0.88</td>
<td>0.80</td>
</tr>
<tr>
<td>B&amp;H - RS</td>
<td>6.26%</td>
<td>0.29</td>
<td>0.18</td>
<td>0.14</td>
<td>0.12</td>
</tr>
<tr>
<td>Difference</td>
<td>-2.6% ns</td>
<td>+0.53***</td>
<td>+3.02***</td>
<td>+0.74***</td>
<td>+0.68***</td>
</tr>
</tbody>
</table>

### Table 2: Fixed percentage SL, with gap: Simulation results

<table>
<thead>
<tr>
<th>Strategy-Model</th>
<th>$\mathbb{E}[r]$</th>
<th>Sharpe</th>
<th>Sortino</th>
<th>RVaR</th>
<th>RES</th>
</tr>
</thead>
<tbody>
<tr>
<td>SL - GEDgap</td>
<td>4.04%</td>
<td>0.81</td>
<td>1.23</td>
<td>0.84</td>
<td>0.69</td>
</tr>
<tr>
<td>B&amp;H - GEDgap</td>
<td>5.01%</td>
<td>0.29</td>
<td>0.13</td>
<td>0.15</td>
<td>0.12</td>
</tr>
<tr>
<td>Difference</td>
<td>-0.97% ns</td>
<td>+0.52***</td>
<td>+1.1***</td>
<td>+0.69***</td>
<td>+0.57***</td>
</tr>
<tr>
<td>SL - ARMAgap</td>
<td>4.06%</td>
<td>0.81</td>
<td>1.21</td>
<td>0.83</td>
<td>0.69</td>
</tr>
<tr>
<td>B&amp;H - ARMAgap</td>
<td>4.93%</td>
<td>0.30</td>
<td>0.13</td>
<td>0.16</td>
<td>0.13</td>
</tr>
<tr>
<td>Difference</td>
<td>-0.87% ns</td>
<td>+0.51***</td>
<td>+1.08***</td>
<td>+0.67***</td>
<td>+0.56***</td>
</tr>
<tr>
<td>SL - RSgap</td>
<td>3.99%</td>
<td>0.80</td>
<td>1.02</td>
<td>0.84</td>
<td>0.72</td>
</tr>
<tr>
<td>B&amp;H - RSgap</td>
<td>5.93%</td>
<td>0.26</td>
<td>0.13</td>
<td>0.11</td>
<td>0.09</td>
</tr>
<tr>
<td>Difference</td>
<td>-1.94% ns</td>
<td>+0.54***</td>
<td>+0.89***</td>
<td>+0.73***</td>
<td>+0.63***</td>
</tr>
</tbody>
</table>

### Table 3: Fixed percentage SL, with gap and flash crash: Simulation results

<table>
<thead>
<tr>
<th>Strategy-Model</th>
<th>$\mathbb{E}[r]$</th>
<th>Sharpe</th>
<th>Sortino</th>
<th>RVaR</th>
<th>RES</th>
</tr>
</thead>
<tbody>
<tr>
<td>SL - GEDcrash</td>
<td>2.34%</td>
<td>0.41</td>
<td>0.07</td>
<td>0.57</td>
<td>0.12</td>
</tr>
<tr>
<td>B&amp;H - GEDcrash</td>
<td>5.03%</td>
<td>0.29</td>
<td>0.12</td>
<td>0.15</td>
<td>0.12</td>
</tr>
<tr>
<td>Difference</td>
<td>-2.69% *</td>
<td>+0.12*</td>
<td>-0.05 ns</td>
<td>+0.42***</td>
<td>+0.00 ns</td>
</tr>
<tr>
<td>SL - ARMACrash</td>
<td>4.27%</td>
<td>0.39</td>
<td>0.07</td>
<td>0.55</td>
<td>0.11</td>
</tr>
<tr>
<td>B&amp;H - ARMACrash</td>
<td>4.90%</td>
<td>0.39</td>
<td>0.13</td>
<td>0.16</td>
<td>0.12</td>
</tr>
<tr>
<td>Difference</td>
<td>-2.63% ns</td>
<td>+0.1 ns</td>
<td>-0.06 ns</td>
<td>+0.39***</td>
<td>-0.01 ns</td>
</tr>
<tr>
<td>SL - RScrash</td>
<td>2.70%</td>
<td>0.52</td>
<td>0.07</td>
<td>0.74</td>
<td>0.18</td>
</tr>
<tr>
<td>B&amp;H - RScrash</td>
<td>5.54%</td>
<td>0.24</td>
<td>0.11</td>
<td>0.10</td>
<td>0.09</td>
</tr>
<tr>
<td>Difference</td>
<td>-2.84% *</td>
<td>+0.28***</td>
<td>-0.04 ns</td>
<td>+0.64***</td>
<td>+0.09*</td>
</tr>
</tbody>
</table>

### Table 4: MA crossover-based SL, with gap and flash crash: Simulation results

<table>
<thead>
<tr>
<th>Strategy-Model</th>
<th>$\mathbb{E}[r]$</th>
<th>Sharpe</th>
<th>Sortino</th>
<th>RVaR</th>
<th>RES</th>
</tr>
</thead>
<tbody>
<tr>
<td>SL - GEDcrash</td>
<td>4.27%</td>
<td>0.31</td>
<td>0.11</td>
<td>0.16</td>
<td>0.13</td>
</tr>
<tr>
<td>B&amp;H - GEDcrash</td>
<td>4.95%</td>
<td>0.29</td>
<td>0.12</td>
<td>0.15</td>
<td>0.12</td>
</tr>
<tr>
<td>Difference</td>
<td>-0.68% ns</td>
<td>+0.02 ns</td>
<td>-0.01 ns</td>
<td>+0.01 ns</td>
<td>+0.01 ns</td>
</tr>
<tr>
<td>SL - ARMACrash</td>
<td>4.24%</td>
<td>0.32</td>
<td>0.10</td>
<td>0.17</td>
<td>0.13</td>
</tr>
<tr>
<td>B&amp;H - ARMACrash</td>
<td>4.96%</td>
<td>0.30</td>
<td>0.12</td>
<td>0.16</td>
<td>0.13</td>
</tr>
<tr>
<td>Difference</td>
<td>-0.72% ns</td>
<td>+0.02 ns</td>
<td>-0.02 ns</td>
<td>+0.01 ns</td>
<td>+0.00 ns</td>
</tr>
<tr>
<td>SL - RScrash</td>
<td>4.74%</td>
<td>0.26</td>
<td>0.10</td>
<td>0.12</td>
<td>0.09</td>
</tr>
<tr>
<td>B&amp;H - RScrash</td>
<td>5.62%</td>
<td>0.25</td>
<td>0.11</td>
<td>0.11</td>
<td>0.09</td>
</tr>
<tr>
<td>Difference</td>
<td>-0.88% ns</td>
<td>+0.01 ns</td>
<td>-0.01 ns</td>
<td>+0.01 ns</td>
<td>+0.00 ns</td>
</tr>
</tbody>
</table>
Hierarchical adaptive sparse grids for option pricing under the rough Bergomi model

Christian Bayer  
Weierstrass Institute for Applied Analysis and Stochastics (WIAS), Berlin, Germany.

Chiheb Ben Hammouda  
King Abdullah University of Science and Technology (KAUST), Thuwal 23955 – 6900, Saudi Arabia  
Email: chiheb.benhammouda@kaust.edu.sa

Raul Tempone*  
King Abdullah University of Science and Technology (KAUST), Thuwal 23955 – 6900, Saudi Arabia  
Email: raul.tempone@kaust.edu.sa

Abstract. The rough Bergomi (rBergomi) model, introduced recently in [1], is a promising rough volatility model in quantitative finance. This new model exhibits consistent results with the empirical fact of implied volatility surfaces being essentially time-invariant. This model also has the ability to capture the term structure of skew observed in equity markets. In the absence of analytical European option pricing methods for the model, and due to the non-Markovian nature of the fractional driver, the prevalent option is to use Monte Carlo (MC) simulation for pricing. Despite recent advances in the MC method in this context, pricing under the rBergomi model is still a time-consuming task. To overcome this issue, we design a novel, alternative, hierarchical approach, based on adaptive sparse grids quadrature, specifically using the same construction as multi-index stochastic collocation (MISC) [2], coupled with Brownian bridge construction and Richardson extrapolation. By uncovering the available regularity, our hierarchical method demonstrates substantial computational gains with respect to the standard MC method, when reaching a sufficiently small error tolerance in the price estimates across different parameter constellations, even for very small values of the Hurst parameter. Our work opens a new research direction in this field, i.e. to investigate the performance of methods other than Monte Carlo for pricing and calibrating under the rBergomi model.

References


*Alexander von Humboldt Professor in Mathematics for Uncertainty Quantification, RWTH Aachen University, Germany.
Optimal life-contingent insurance under bid-ask spreads

NIKOLAY BONDARENKO
MACQUARIE UNIVERSITY, 4 EASTERN ROAD, NORTH RYDE, SYDNEY, NSW 2109, AUSTRALIA
Email: nikolay.bondarenko@hdr.mq.edu.au

GEOFFREY KINGSTON
MACQUARIE UNIVERSITY, 4 EASTERN ROAD, NORTH RYDE, SYDNEY, NSW 2109, AUSTRALIA
Email: geoff.kingston@mq.edu.au

PAVEL SHEVCHENKO
MACQUARIE UNIVERSITY, 4 EASTERN ROAD, NORTH RYDE, SYDNEY, NSW 2109, AUSTRALIA
Email: pavel.shevchenko@mq.edu.au

Abstract. We re-examine optimal insurance in a Merton-type setting when life insurance and/or life annuities carry loads. Previous contributions incorporate a parameter corresponding to the bid price of life insurance but overlook incorporating a parameter corresponding to the ask price; we rectify this omission. If the bid-ask spread is positive, nonparticipation in the market for life-contingent insurance is optimal during two positive intervals, one in midlife and the other at the end of life’s maximal span.

1. Introduction. We re-examine optimal insurance when life insurance and/or life annuities carry loads. Richard (1975) and Pliska and Ye (2007) are the two classic contributions to this line of research. They extend the Merton portfolio model of an investor’s life cycle by exploiting the insight that a long position in life insurance is effectively a short position in life annuities, and vice versa.

The Richard-Pliska-Ye model implies that optimal participation in the market for life-contingent insurance is continuous (i.e. lifelong) even if insurance loads are heavy and the investor is close to being risk neutral. We find instead that a positive bid-ask spread induces two intervals during which the investor sits on the sidelines of the market for life-contingent insurance, one in midlife and the other at the end of life’s maximal span. The intuition for the midlife nonparticipation interval is that self-insurance is a better deal for a time than either life insurance or life annuities. In the case of the late-life interval, once you confidently expect to live only for a few more years you become increasingly less inclined to buy a loaded life annuity to avoid the risk of outliving your resources, assuming you have a bequest motive.

2. Model. Following Richard (1975) and Pliska and Ye (2007), among others, we assume a fixed planning horizon $T$, which can be interpreted as the maximum possible span over which an investor could conceivably consume and earn a wage. An investor alive at time $t$ has a random lifespan $\tau$, $0 \leq \tau < \infty$, with a probability density function $f(t)$. The investor’s wage is a general deterministic process $y(t) \geq 0$, $0 \leq t \leq T$. To highlight the implications of introducing an ask price for life insurance, we abstract from risky investments, in line with Pliska and Ye (2007) but departing from Richard (1975), Ye (2008), Aase (2017) and many other contributions. The investor’s wealth constraint is

$$dx(t) = rx(t) dt - c(t) dt - p(t) dt + y(t) dt$$

(1)

where the new variables are financial wealth $x(t)$, consumption $c(t)$, and wage $y(t)$, where $t \leq \min[\tau, T]$. The investor’s legacy when she dies at time $t = \tau$ is

$$z(t) = x(t) + \frac{p(t)}{\eta(t)} = \begin{cases} x(t) \\ x(t) + \frac{p(t)}{\phi(t)} \end{cases}$$

(2)
Optimal life-contingent insurance under bid-ask spreads

according to whether she is long, not invested or short in life insurance.

For simplicity, and following part of Richard (1975), most of Pliska and Ye (2007), and many other contributions, we assume constant relative risk aversion (the penultimate section specializes further, to log utility). Maximum expected utility over the set of admissible consumption and insurance policies is

\[ V(x, 0) \equiv \sup_{c, p} \mathbb{E}_{0, x(0)} \left[ \int_0^{T_\Lambda} e^{-\rho t} \frac{c(t)}{\gamma} dt + e^{-\rho \tau} \frac{z(t)}{\gamma} 1_{\{\tau \leq T\}} + e^{-\rho T} \frac{x(T)}{\gamma} 1_{\{\tau > T\}} \right] \]  

(3)

where the new variables are the expectations operator \( \mathbb{E} \), utility-function curvature parameter \( \gamma < 1 \), and rate of time preference \( \rho > 0 \). By a standard argument, the value function can be restated as one with a non-random terminal time \( T \):

\[ J(x(t); c, p) = \int_t^T \left[ f(s, t) e^{-\rho s} \frac{z(s)}{\gamma} + \bar{F}(s, t) e^{-\rho s} \frac{c(s)}{\gamma} \right] ds + \bar{F}(T, t) e^{-\rho T} \frac{x(T)}{\gamma} , \]  

(4)

where \( f(s, t) \) is the probability density for death at time \( s \) conditional upon surviving at time \( t \leq s \), and \( \bar{F}(s, t) \) is the probability for survival at time \( s \) conditional upon survival at time \( t \leq s \). These probabilities are related to \( \lambda(t) \) by

\[
\begin{aligned}
    f(s, t) &= \frac{f(s)}{F(t)} = \lambda(s) \exp \left\{ - \int_t^T \lambda(u) \, du \right\} , \\
    \bar{F}(s, t) &= \frac{\bar{F}(s)}{F(t)} = \exp \left\{ - \int_t^s \lambda(u) \, du \right\} .
\end{aligned}
\]  

(5)

We solve this model piecewise, using dynamic programming to obtain solutions for the insured intervals. However, that method appears not to work for the nonparticipation intervals, so we turn to the martingale method for them.

References


Understanding generalisation in noisy time series forecasting

ANASTASIA BOROVYKH
CWI, AMSTERDAM, THE NETHERLANDS
EMAIL: borovyck.a@hotmail.com

Abstract. In this presentation we study the loss surface of neural networks for noisy time series forecasting. We aim to gain insight into the effects of deep versus wide networks on the loss surface and the structure of the critical points. In extrapolation problems for noisy time series, neural networks, due to their overparametrization, tend to overfit and the behavior of the model on the training data does not measure accurately the behaviour on unseen data due to e.g. changing underlying factors in the time series. Avoiding overfitting and finding a pattern in the data that persists for a longer period of time can thus be very challenging. In this talk we quantify what the neural network has learned using the structure of the loss surface of multi-layer neural networks. We study the role that the optimization method plays in the generalisation capabilities and gain insight into which minima are able to generalise well based on the spectrum of the Hessian matrix and the smoothness of the learned function with respect to the input. Based on the structure of the time series we quantify how to make the trade-off between the complexity of the learned function and the ability of the function to fit the data well so that the network can generalise well.
A Reinforced Urn Process Modeling of Recovery Rates and Recovery Times

Dan Cheng  
Applied Probability Group, Delft University of Technology, The Netherlands  
Email: d.cheng@tudelft.nl

Pasquale Cirillo  
Applied Probability Group, Delft University of Technology, The Netherlands  
Email: p.cirillo@tudelft.nl

Abstract. Answering a major demand in modern credit risk management, we propose a Bayesian nonparametric survival approach for the modeling of the recovery rates and the recovery times of a portfolio of defaulted exposures, whose recovery trajectories are assumed exchangeable [1].

The model we introduce is called Recovery Reinforced Urn Process (R-RUP), and it represents a special type of a more general class of combinatorial stochastic processes [2]. A R-RUP \( \{X_n\}_{n \geq 0} \) is characterized by the following elements:

- A countable state space \( S = N_0 \times \{0, 1, ..., m\} \), containing all the couples \( (t, l) \) of recovery times \( t \) and recovery levels \( l \) the process \( \{X_n\} \) can visit with positive probability. Levels \( 0 \) to \( m - 1 \) are a discretization of the recovery rates from 0\% to 100\%, while level \( m \) represents the termination level.

- On every element \( s \in S \) is centered a Pólya urn [3], containing balls of colors belonging to a set \( C = \{c_0, c_1, ..., c_m\} \). Each color from \( c_0 \) to \( c_m \) corresponds to a given recovery level, as if we color them. Every urn can have a different composition, with some colors missing, but we assume that no empty urn exists.

- A law of motion \( q: S \times C \to S \) indicating how the sampling of the different urns drives the process \( \{X_n\} \). In particular, we assume:
  - \( q((t, l), c_i) = (t + 1, l) \), for \( i \leq l \leq m \). In words: if, from the urn centered in \( (t, l) \), we extract a ball whose color corresponds to level \( l \) or lower, the process moves to the next time unit \( t + 1 \), but stays at the same recovery level \( l \). From \( (t, l) \) we move to \( (t + 1, l) \).
  - \( q((t, l), c_i) = (0, i) \), for \( l < i \leq m \). The process jumps to level \( i > l \) while time is reset to 0. Resetting time at every jump helps in counting the time units the R-RUP spends in each recovery level it visits.
  - \( q((1, m), c_m) = (0, 0) \). The process \( \{X_n\} \) can stay in level \( m \) only for one time unit, after which it reaches \( (1, m) \) and it is reset to \( (0, 0) \). This last movement allows us to define a recurrent R-RUP, which fully exploits the reinforcement mechanism of Pólya urns (and the embedded learning).

Without any loss of generality, fix an initial state \( s_0 \). A R-RUP \( \{X_n\} \) on \( S \) with initial state \( s_0 = (0, 0) \) is defined recursively as follows: set \( X_0 = s_0 \), and for all \( n \geq 1 \), if \( X_{n-1} = s_{n-1} \in S \), sample a ball from the urn \( U(s_{n-1}) \) associated with \( s_{n-1} \). Now, register the color of the ball, say \( c \in C \), put it back in \( U(s_{n-1}) \), and Pólya-reinforce the urn with \( r > 0 \) balls of the same color. This increases the probability of picking again that color in the future, if the urn is sampled again: the higher \( r \) the stronger the update. Finally, using the rule of motion \( q \), set \( X_n = q(s_{n-1}, c) \). The sequence \( \{X_n\} \) is a recovery reinforced urn process with initial state \( s_0 \) and reinforcement \( r \). A R-RUP is thus a reinforced random walk on a state space of Pólya urns.

It can be shown that the R-RUP generates a mixture of semi-Markov chains, whose mixing (de Finetti) measure is a special random distribution, named beta-Stacy Dirichlet. We study the properties of such a process, showing for instance its conjugacy, something that proves extremely useful in applications, and in Bayesian prediction.

Our model has several advantages. First, it allows for the elicitation and exploitation of prior knowledge and experts’ judgements (via the urns’ compositions), and for the constant update of this
information over time, as soon as new data become available (via the Pólya-reinforcement). The possibility of using an a priori is a great advantage in a field, credit risk management, in which there are often data problems, which only the knowledge and intuition of trained experts can try to compensate. Second, the model is able to deal with right-censored observations, a common problem in recovery modeling. Third, being a Bayesian model, our construction easily allows for prediction about new observations and trajectories, given the available information and the a priori. Finally, the semi-Markovianity embedded in our model guarantees more flexibility in applications, with respect to the usual approaches in the literature.

References


A urn-based nonparametric modeling of the dependence between PD and LGD with an application to mortgages

PASQUALE CIRILLO
APPLIED PROBABILITY GROUP, DIAM, DELFT UNIVERSITY OF TECHNOLOGY, THE NETHERLANDS
EMAIL: P.Cirillo@tudelft.nl
URL: http://www.pasqualecirillo.eu

DAN CHENG
APPLIED PROBABILITY GROUP, DIAM, DELFT UNIVERSITY OF TECHNOLOGY, THE NETHERLANDS
EMAIL: D.Cheng@tudelft.nl

Abstract. We propose an alternative approach to the modeling of the positive dependence between the probability of default and the loss given default in a portfolio of exposures, using a bivariate urn process. The model combines the power of Bayesian nonparametrics and statistical learning, allowing for the elicitation and the exploitation of experts’ judgements, and for the constant update of this information over time, every time new data is available.

A real-world application on mortgages is described using the Single Family Loan-Level Dataset by FreddieMac.

1. Introduction.

The ambition of this paper is to present a new way of modeling the empirically-verified positive dependence between the Probability of Default (PD) and the Loss Given Default (LGD) using a Bayesian nonparametric approach, based on urns and beta-Stacy processes [19]. The model is able to learn from the data and it improves its performances over time, compatibly with the machine learning paradigm. Similarly to the recent construction of [8] for recovery times, this learning ability is mainly due to the underlying urn model.

The PD and the LGD are two fundamental quantities in modern credit risk management. The PD of a counterparty indicates the likelihood that that counterparty will default and thus not be able to fulfil its debt obligations. The LGD represents the percent loss, in terms of the notional value of the exposure, known as the Exposure-at-Default or EAD, one actually experiences when a counterparty defaults and every possible recovery process is over. Within the Basel framework [4, 6], a set of international standards developed by the Basel Committee on Banking Supervision to harmonize the banking sector and improve the way banks manage risk, the PD and the LGD are considered pivotal risk parameters for the quantification of the minimum capital requirements for credit risk. In particular, under the so-called Internal Rating-Based (IRB) approaches, both the PD and the LGD are inputs of the main formulas for the computation of the risk-weighted assets [5].

Surprisingly, in most theoretical models, and most of all in the formulas suggested by the BCBS, the PD and the LGD are assumed to be independent, even though several empirical studies have shown that there is a non-negligible positive dependence between them [1, 16, 21]. Borrowing from the terminology developed in the field of credit valuation adjustment (CVA) [14], this dependence is often referred to as wrong-way risk (WWR). Simply put, WWR is the risk that the possible loss generated by a counterparty increases with the deterioration of its creditworthiness. Ignoring this WWR can easily lead to an unreliable estimation of credit risk for a given counterparty [21].

The link between the probability of default and the loss given default is particularly important when dealing with mortgages, as shown by the 2007-2008 financial crisis, which was triggered by an avalanche of defaults in the US subprime mortgage market [12], with a consequent drop in estate prices, and thus in the recovery rates of the defaulted exposures [14]. The financial crisis was therefore one of the main drivers for the rising interest in the joint modeling of PD and LGD, something long ignored both in the academia and in the practice—including regulators.
Our contribution to this open problem is represented by the present paper, which finds its roots in the recent literature about the use of urn models in credit risk management [3, 9, 18, 8]. A real-world application in the field of mortgages is provided.

2. Model. The model we propose is based on powerful tools of Bayesian nonparametrics, like the beta-Stacy process of [19] and Reinforced Urn Processes (RUP) [17].

Consider a portfolio \( P \) containing \( m \) exposures. For \( i = 1, \ldots, m \), let \( X_i \) and \( Y_i \) represent the PD, respectively the LGD, of the \( i \)-th counterparty, when discretised and transformed into levels, in a way similar to what [8] propose in their work. In other terms, one can split the PD and the LGD into \( l = 0, \ldots, L \) levels, such that for example \( l = 0 \) indicates a PD or LGD of 0%, \( l_1 \) something between 0% and 5%, \( l = 2 \) a quantity in (5%, 17%), and so on until the last level \( L \). The levels do not need to correspond to equally spaced intervals, and this gives flexibility to the modeling. Clearly, the larger \( L \) the finer the partition we obtain (more details in the full paper).

Let \( \{A_i\}_{i=1}^m \), \( \{B_i\}_{i=1}^m \) and \( \{C_i\}_{i=1}^m \) be three independent LS sequences generated by three independent two-color RUPs \( \{Z^A_i\}_{j \geq 0}, \{Z^B_i\}_{j \geq 0}, \{Z^C_i\}_{j \geq 0} \). As we know, the sequence \( \{A_i\}_{i=1}^m \) is exchangeable, and its de Finetti measure is a beta-Stacy process \( F_A \) of parameters \( \{\alpha^A_j, \beta^A_j\}_{j \in \mathbb{N}_0} \). Similarly, for \( \{B_i\}_{i=1}^m \) and \( \{C_i\}_{i=1}^m \) we have \( F_B \) with \( \{\alpha^B_j, \beta^B_j\}_{j \in \mathbb{N}_0} \), and \( F_C \) with \( \{\alpha^C_j, \beta^C_j\}_{j \in \mathbb{N}_0} \).

Now, following [7], let us assume that, for each exposure \( i = 1, \ldots, m \), we have

\[
X_i = A_i + B_i, \\
Y_i = A_i + C_i.
\]

This construction builds a special dependence between the discretised PD and the LGD of our exposures. In fact, we are assuming that for each counterparty \( i \) there exists a common factor \( A_i \), influencing both the PD and the LGD, while \( B_i \) and \( C_i \) can be seen as idiosyncratic components (conditionally on \( A_i \), \( X_i \) and \( Y_i \) are clearly independent).

Since we can write \( Y_i = X_i - B_i + C_i \), it is clear that we are assuming a linear dependence between \( X_i \) and \( Y_i \). This is compatible with several empirical findings, like [2] or [16].

Furthermore, given Equation (1) and the properties of the sequences \( \{A_i\}_{i=1}^m \), \( \{B_i\}_{i=1}^m \) and \( \{C_i\}_{i=1}^m \), we can immediately observe that

\[
\text{Cov}(X_1, Y_1) = \text{Var}(A_1) \geq 0, \\
\text{Cov}(X_{m+1}, Y_{m+1}|A_m, B_m, C_m) = \text{Var}(A_{m+1}|A_m) \geq 0, \quad \text{for } m \geq 2,
\]

where \( \text{Var} \) is the variance, \( \text{Cov} \) the covariance, \( A_m = [A_1, \ldots, A_m] \), and similarly \( B_m, C_m \). Therefore, with the bivariate urn model we can only model positive dependence. Again, this is totally in line with our analysis—the study of wrong-way risk that by definition is a positive dependence between PD and LGD—and with the empirical literature, as discussed in Section 1. However, it is important to stress that the model in Equation (1) cannot be used when negative dependence is possible.

Always from Equation (1), we can verify that the sequence \( \{(X_i, Y_i)\}_{i=1}^m \) is exchangeable. An implicit assumption of our model is therefore that the \( m \) counterparties in \( P \) are exchangeable. As observed in [15], exchangeability is a common assumption in credit risk, as it is seen as a relaxation of the stronger hypothesis of independence (think about Bernoulli mixtures and the beta-binomial model). All in all, what we ask is that the order in which we observe our counterparties is irrelevant to study the joint distribution of their PDs and LGDs, which is therefore immune to changes in the order of appearance of each exposure. Exchangeability and the fact that \( X_i \) and \( Y_i \) are conditionally independent given \( A_i \) suggest that the methodology we are proposing falls under the larger umbrella of mixture models [15, 11].

Assume now that we have observed \( m \) exposures, and we have registered their actual PD and LGD, which we have discretised to get \( \{(X_i, Y_i)\}_{i=1}^m \). The construction of Equation (1), together with the properties of the beta-Stacy processes involved, allows for a nice derivation of the predictive distribution for a new exposure \((X_{m+1}, Y_{m+1})\), given the observed couples \((X_m = x_m, Y_m = y_m)\). This can be extremely useful in applications, when one is interested in making inference about the PD, the LGD and their relation.
In fact
\[ P[X_{m+1} = x, Y_{m+1} = y | X_m = x_m, Y_m = y_m] = P[X_{m+1} = x, Y_{m+1} = y, X_m = x_m, Y_m = y_m] \]
\[ = P[X_{m+1} = x, Y_{m+1} = y | X_m = x_m, Y_m = y_m]. \tag{3} \]

From a theoretical point of view, computing Equation (3) just requires to count the number balls in the urns behind \( \{A_i\}_{i=1}^m \), \( \{B_i\}_{i=1}^m \) and \( \{C_i\}_{i=1}^m \) and use the common predictive equations for urn processes [10]; something that for a small portfolio can be done explicitly. However, when \( m \) is large, it becomes numerically unfeasible to perform all those sums and products.

Luckily, developing an alternative Markov Chain Monte Carlo algorithm is simple and effective. It is sufficient to go through the following steps.

1. Given the observations \( X_m = x_m \) and \( Y_m = y_m \), the sequence \( A_m = (A_1, \ldots, A_m) \) is generated via a Gibbs sampling. The full conditional of \( A_m \), \( P[A_m = a_m | A_{m-1} = a_{m-1}, X_m = x_m, Y_m = y_m] \) is such that

\[
P[A_m = a_m | A_{m-1} = a_{m-1}, X_m = x_m, Y_m = y_m] \propto P[A_m = a_m | A_{m-1} = a_{m-1}] \times P[X_m - A_m = x_m - a_m | B_{m-1} = b_{m-1}] \times P[Y_m - A_m = y_m - a_m | C_{m-1} = c_{m-1}].
\]

Since \( \{A_j\}_{j=1}^m \) is exchangeable, all the other full conditionals, \( P[A_j = a_j | A_{-j} = a_{-j}, X_m = x_m, Y_m = y_m] \), have an analogous form.

2. Once \( A_m \) is obtained, compute \( B_m = X_m - A_m \) and \( C_m = Y_m - A_m \).

3. The quantities \( A_{m+1}, B_{m+1}, \) and \( C_{m+1} \) are then sampled according to their beta-Stacy predictive distributions \( P(A_{n+1} | A_m) \), \( P(B_{n+1} | B_m) \), and \( P(C_{n+1} | C_m) \) using the properties of the underlying RUPs.

4. Finally, set \( X_{m+1} = A_{m+1} + B_{m+1} \) and \( Y_{m+1} = A_{m+1} + C_{m+1} \).

3. Results. The performances of the bivariate urn model are tested on data coming from the Single Family Loan-Level Dataset Sample by Freddie Mac [13].

As an example, Figure 1 shows, for the worst FICO score group in the data, the very good fitting performances of the model for the marginal distributions of the discretised PD and LGD respectively. Each subfigure shows the elicited prior, the empirical cumulative distribution function (ecdf) and the posterior, as obtained via learning and reinforcement. In the figure shown, discretisation is obtained by simple rounding, while the priors are discrete uniforms. A two sample Kolmogorov-Smirnov (KS) test does not reject the null hypothesis of same distribution between the ecdfs and the relative posteriors (p-values stably above 0.05).

Figure 2 shows the bivariate distributions we obtain for the discretised PD and LGD, in case of weak and strong Poisson a priori, which we assume to be the result of experts’ judgement.

Regarding the numbers, the correlation is properly captured, as well as the mean and the variances of the marginals, especially when the strength of beliefs is smaller. More details are available in the full paper, which will be discussed at the conference.

References


Urn modeling of the PD/LGD dependence

(a) PD
(b) LGD

Figure 1: Prior, posterior and ecdf for the discretised PD and LGD in the "Very poor" FICO rating class.

(a) Weak Prior
(b) Strong Prior

Figure 2: Bivariate density distribution of PD and LGD in the "Very poor" FICO rating class, with Poisson priors and strength of belief equal to 1.


A spectral method for an optimal investment problem with transaction costs under potential utility

JAVIER DE FRUTOS*  
INSTITUTO DE MATEMÁTICAS (IMUVA), UNIVERSIDAD DE VALLADOLID, PASEO DE BELÉN 7, VALLADOLID, SPAIN  
EMAIL: frutos@mac.uva.es

VÍCTOR GATÓN  
INSTITUTO DE MATEMÁTICAS (IMUVA), UNIVERSIDAD DE VALLADOLID, PASEO DE BELÉN 7, VALLADOLID, SPAIN  
EMAIL: vgaton@mac.uva.es

Abstract. This talk concerns the numerical solution of the finite-horizon Optimal Investment problem with transaction costs under Potential Utility (see [2]). The problem is initially posed in terms of evolutive HJB equation with gradient constraints and in [1] the problem is reformulated as a non-linear parabolic double obstacle problem posed in one spatial variable and defined in an unbounded domain.

The restatement of the problem in polar coordinates allows to pose the problem in one spatial variable in a finite domain, avoiding many technical difficulties of the previous statement of the problem. If high precision is required, the spectral numerical method proposed becomes more efficient than simpler methods as finite differences for example.

Current work on the problem includes the extension of the numerical method to deal with multiple stocks.

Keywords: Optimal Investment, Potential Utility, Transaction costs, Spectral method.

References


*Research supported by MINECO/AEI proyect MTM2016-78995-P and by Junta de Castilla y Len VA105G18 and VA024P17 cofinanced by FEDER funds (EU).
Portfolio risk and the quantum majorization of correlation matrices

ANDREA FONTANARI
DELFT UNIVERSITY OF TECHNOLOGY
EMAIL: a.fontanari@tudelft.nl

IDDO ELIAZAR
INDEPENDENT RESEARCHER

PASQUALE CIRILLO
DELFT UNIVERSITY OF TECHNOLOGY

CORNELIS W. OOSTERLEE
DELFT UNIVERSITY OF TECHNOLOGY AND CWI

Abstract. We propose quantum majorization as a way of comparing and ranking correlation matrices, with the aim of assessing portfolio risk in a unified framework. Quantum majorization is a partial order in the space of correlation matrices, which are evaluated through their spectra.

We discuss the connections between quantum majorization and an important class of risk functionals, and we define two new risk measures able to capture interesting characteristics of portfolio risk.

1. Extended Abstract. Consider a portfolio $P$ containing $n$ assets. A common way to deal with the dependence structure of $P$—and thus with portfolio risk—is to consider its $n \times n$ correlation matrix $C$. We refer to the correlation matrix in the most general sense, as a real, symmetric, positive semidefinite, Hermitian matrix, whose entries are correlation coefficients according to some definition, usually Pearson’s one, all lying in the interval $[-1, 1]$, with all ones on the main diagonal.

The main idea of this work is to find a way of comparing correlation matrices, and of extracting the portfolio risk information they contain, so that they can be ranked. The comparison can be over time, if one studies the evolution of the correlation matrix for a given portfolio, but it can also be cross-sectional, comparing different portfolios at the same time, for example to look for the one minimizing portfolio risk. The only requirement is that the size of the matrices is the same, i.e. only portfolios having the same number of assets are considered.

Our proposal is to use an ordering, developed in the field of quantum mechanics [1, 3], called quantum majorization, to study the dynamics of the entropy of a quantum system, applying it to correlation matrices. To the best of our knowledge, this is the first time such an ordering is used in finance.

Quantum majorization is a partial order to rank matrices looking at their eigenvalues. The use of eigensystems to study multivariate dependence is not at all new [4], but we show how the spectrum of a correlation matrix can be used to capture relevant features of portfolio risk in a brand new way.

We first use quantum majorization to evaluate the goodness of existing risk measures, such as the portfolio total variance, as well as build a framework which allows for the specification of entire new families of risk measures. In this way we introduce the $M_\lambda$ class of risk functionals, i.e. functions isotonic to quantum majorization, and whose aim is to capture the (monotonic) dependence embedded into portfolio correlation matrices. An important property of such a class, is that under quantum majorization, all risk functionals in $M_\lambda$ are comonotonic. The implication is that, if we are able to identify majorization, then the choice of the risk functional becomes secondary, as they will all behave in the same way, indicating an increase (or a decrease) of portfolio risk. It is when the ordering does not hold—as we shall see—that risk functionals may give inconsistent information.

With respect to single risk measures, quantum majorization thus provides a stronger characterization of risk and dependence among correlation matrices. We are therefore able to provide a unifying approach.
to the analysis of portfolio risk and correlation: we introduce several tools, we discuss their properties, and we show how to use them in practice. In doing so, we will avoid all unnecessary sophistication, giving space to financial interpretability and usability.

Secondly, we exploit the connection between partial orders and directed acyclical graphs, DAG, to build a network representation of portfolios where a directed edge is placed between nodes, the portfolios, if the quantum majorization relation is present. Figure 1 provides a pictorial representation of the so constructed network using the 30 components of the Industrial Dow Jones. Using this interpretation we were able to build an alarm system based on graphical features via a spectral embedding algorithm and a k-mean clustering. This approach is novel to risk literature despite graphical features for price prediction are being used more and more in investing strategies.

In conclusion, we develop a mathematical machinery to extract new information from the market and use this new information to assess risk measures and build warning systems based on the dynamic of portfolio dependence.

Finally, we argue that our approach can be also extended to trading strategies as a way to generate a trading signal depending on the presence and persistence of quantum majorization in the stock markets, applications may be found for example in factor investing as well as basket option trading.

References


Empirics on CPPI Design Risk

RAQUEL M. GASPAR*
ISEG and REM - CEMAPRE, UNIVERSIDADE DE LISBOA, PORTUGAL
EMAIL: Rmgaspar@iseg.ulisboa.pt

Abstract. This study aims at empirically analyse design risk associated with some portfolio insurance strategies. Although the focus is on constant proportion portfolio insurance (CPPI) structures, we also look at other portfolio insurance strategies.

The paper is based upon real market data on several markets indices representing a variety of different risky assets, both from stock and corporate bond markets. These indices are considered as the underlying risky assets.

The empirical approach of this study is different from what has been done in most of the literature as we make no assumption on a particular model for the underlying risky asset dynamics, while most studies assume either a geometric brownian motion or a more general model that include jumps. Our results, are thus, model free and based upon real observed data.

Overall, this study strengths the idea that CPPI strategies suffers from a serious design problem, making it extremely path-dependent and uninteresting to almost all investors.

1. CPPIs and Portfolio Insurance. The idea of introducing insurance in investment portfolios was first proposed by Leland and Rubinstein [25]. The main motivation was to prevent the contagious disinvestment movements observed in the stock market crash of 1973-74, which led to the loss of significant potential gains in the subsequent 1975 rise.

A portfolio insurance (PI) strategy consists of an asset allocation strategy between a risk-free asset and a risky asset, so that the combination would give the investor both security, as a percentage of the initial investment guaranteed at maturity, and the possibility of some participation in upside performance of the risky asset.

Leland and Rubinstein [25] developed the first PI strategy, the Option Based Portfolio Insurance (OBPI), realising that the risky asset can be insured by a put option written on it and whose strike price is the amount to be insured. Although listed options are most of the times not available for long maturities, this difficulty could be easily replaced by theoretically modeling the risky asset dynamics. Following the work of Merton [26], at the time the obvious was to consider the recently developed Black and Scholes [11] pricing model (BS model). Using the BS model, Rubinstein [30] proposed an alternative to the static OBPI, based upon t dynamic option replication. This synthetic OBPI, is an asset allocation strategy between a risky and a risk-free asset, based upon the delta-hegaing of options, which is extremely popular in the industry. In fact, in most of the literature term “synthetic OBPI” is shortened to just “OBPI”, as the static OBPI is rarely used.

A few years later, Perold [28] proposed an alternative PI strategy, the Constant Proportion Portfolio Insurance (CPPI). CPPI strategies were understood as a possible solution of the Merton [26] problem, for an investor with hyperbolic absolute risk aversion (HARA) utility function. That is, CPPIs were proposed as possible solutions to a very concrete mathematical problem, under extreme assumptions, not only on the risky asset dynamics, but also on the way investors make decisions. For further discussion of this issue see, for instance, [22].

No matter the reasons underlying its creation, a CPPI strategy can be understood as a model-free dynamic asset allocation between a risky and a risk-free asset, that is able to guarantee a certain percentage of the initial investment at maturity, just like an OBPI. From a design point of view, however, it is considerably simpler than the synthetic OBPI, as its risky exposure relies on a fixed multiplier

*The author was partially supported by the Project CEMAPRE - UID/MULTI/00491/2019 financed by FCT/MCTES through national funds.
Empirics CPPI design risk

(instead of a dynamic delta). Its simpler implementation made CPPIs appealing to a great number issuers, who did not need to rely on any model to manage them.

The term “constant proportion” derives from the fact that for every rebalancing date, the amount of the portfolio invested in the risky asset (exposure) is proportional to the so called “cushion”. This cushion is nothing but the difference between the total strategy value at that instant, and the present value of the amount insured at maturity (the “floor”). The proportionality factor (the —‘multiplier”) is fixed at inception for the entire investment period, and is supposed to be greater than one. Common multipliers in real life products range from three to seven or higher.

Since the first appearance of OBPI and CPPI strategies, an extensive literature has sprouted on the subject, with different objectives and methodologies.

Most studies concerned the theoretical properties of continuous-time PI, see for instance [12] or [9], and references therein. This first stream of the literature focused mainly in solving an optimisation mathematical problem that arises when we assume investors maximize their end-of-period expected utility, as proposed in [31]. Recently, [3] show that even under these classical assumptions, CPPIs would only be optimal if the exogenously given guarantee enters directly into utility maximisation problem itself. Nowadays, however, since the emergence of behavioural finance, the classical expected utility theory itself has been under discussion. For an overview on behavioural finance we refer the reader to survey [4]. Dicht [17] and Gaspar and Silva [18] evaluate to which extent different behavioural theories would help understanding PI investments. Their results show these theories may indeed help understanding the usage of some (but for all) PI strategies. In particular, CPPI strategies cannot be understood in this context.

A second stream of the literature focused on comparing PI strategies with respect to performance, distribution of returns and/or stochastic dominance, using Monte Carlo simulations, and relying on theoretical models for the underlying risky asset. The initial approach considered the BS model – see e.g. [10], [6, 7] and [8]– and compared strategies mainly in terms of risk and performance measures. Later analyses focused on distribution properties and stochastic dominance – see, for instance, [2] and [33], [8] or [16]. The literature looked at alternative models for the underlying risky asset, including jumps-models as in [15] or regime switching models as in [32]. The results are mixed, but recent studies show that CPPI strategies, present an odd distribution of returns with high probability of returns very close to the floor and low probability of extreme positive returns. Moreover, CPPIs seem to be stochastically dominated by naive portfolio strategies.

A third stream of the literature uses observed empirical densities instead of models, which is considered to be more realistic. [13], [23] compare the most used PI strategies – OBPI and CPPI – arriving at the conclusion that CPPIs have better performance only in bear and no-trend markets. [1], however, show that even in that case, CPPIs, get stochastically dominated by naive strategies.

Finally, more recently, a fourth stream has emerged, first identifying problems with the design of CPPI strategies and then proposing modifications to overcome identified problems. In terms of modifications [27] summarized some of the proposals until then, from which we highlight the “cushion insurance” of [29]. After that it is worth mention, the dynamics proportions proposed by [14] or the contingent retracted floor of [24]. [21] proposes alternative strategies with pre-specified distributional properties that present much better results than CPPIs and [5] suggests modifications taking the perspective of the pension funds industry and long-term investors.

For a more detailed overview on PI, we refer to the survey study [19] and the encyclopedia article [20], respectively.

The study here presented is closest to the third stream of the literature, but differs from the existing literature by: (i) Taking a worldwide perspective, considering a wide variety of indices, (ii) Using bootstrapping and sampling technique to determine, (iii) Trying to measure design risk.

2. Empirics. Concretely, we rely on observed daily returns over the past 20 years on the indices – SP500 (SPX), Euro Stoxx 50 (SX5E), MSCI World (MXWO), MSCI Emerging Market (MXEF) and iBoxx EuroCorp TR (QW5A). The criteria was to choose indices that have been used in real life CPPI products, either as the risky underlying or as a proxy to the underlying risk portfolio.
Using bootstrapping techniques we are able to empirically simulate daily returns for the assumed underlying risky asset. Using Monte Carlo simulation on the empirical distribution of the underlying asset, we can not only easily find the empirical payoff distribution of portfolio insurance strategies, but we are also able to compute all relevant statistics, not only at maturity but also during the product’s life span.

We look into CPPI strategies with different multipliers, but also into the classical option based portfolio insurance (OPBI) and some naive strategies such as the stop loss portfolio insurance (SLPI). A typical portfolio insurance strategy provides a capital guarantee $F$ at maturity $T$ and some possible participation in the upside potential of a risky underlying asset, if the underlying performs well.

As expected the value of portfolio insurance strategies depend on: (i) $F$ the level of capital guarantee, (ii) $T$ the maturity, or (iii) the value at maturity of the underlying risky asset $S_T$ or, if path dependent, on the actual evolution of the underlying risky asset, $\{S_t, 0 \leq t \leq T\}$. However in the case of CPPIs, the structure’s performance depends strongly (iv) on the multiplier $m$; (v) the rebalancing frequency, (vi) any possible deductions from the investment (namely fees and/or coupons). We call the dependence of CPPI structures on these variables (iv)-(vi) which are unrelated to the underlying risky asset: design risk.

References


Collocating Local Volatility: A Competitive Alternative to Stochastic Local Volatility Models

LECH A. GRZELAK*
DELFT INSTITUTE OF APPLIED MATHEMATICS, DELFT UNIVERSITY OF TECHNOLOGY,
DELFT, THE NETHERLANDS
Email: L.A.Grzelak@tudelft.nl

ANTON v.d. STOEP
PRICING MODEL VALIDATION, RABOBANK, UTRECHT, THE NETHERLANDS
Email: Anton.van.der.Stoep@rabobank.com

KEES OOSTERLEE
CWI, NATIONAL RESEARCH INSTITUTE FOR MATHEMATICS AND COMPUTER SCIENCE,
AMSTERDAM, THE NETHERLANDS
Email: C.W.Oosterlee@cwi.nl

Abstract. We discuss a competitive alternative to stochastic local volatility models, namely the Collocating Local Volatility (CLV) model, introduced in [12]. The CLV model consists of two elements, a ‘kernel process’ that can be efficiently evaluated and a local volatility function. The latter, based on stochastic collocation —e.g. [2, 27, 10, 26]— connects the kernel process to the market and allows the CLV model to be perfectly calibrated to European-type options. In this article we consider three different kernel process choices: the Ornstein-Uhlenbeck (OU) and Cox-Ingersoll-Ross (CIR) processes and the Heston model. The kernel process controls the forward smile and allows for an accurate and efficient calibration to exotic options, while the perfect calibration to liquid market quotes is preserved. We confirm this by numerical experiments, in which we calibrate the OU-CLV, CIR-CLV and Heston-CLV models to FX barrier options.

1. Introduction. In the last decade, the class of stochastic local volatility (SLV) models —e.g. [15, 16, 20, 5, 23, 19, 24, 25]— has been given a lot of attention. According to Lipton et al., SLV models are the de facto standard for pricing FX options in practice [17]. They combine the desirable features of the standard Local Volatility model [6, 8] —an almost perfect calibration to liquid European-type options—and well-established stochastic volatility models such as the Heston model [14] and the SABR model [13], which often yield realistic forward smiles and prices of exotics. However, SLV models involve some conditional expectation that is non-trivial and may be expensive to evaluate.

In this article we consider an alternative to SLV models, namely the Collocating Local Volatility (CLV) model, introduced in [12]. The CLV model is composed of a kernel process and a local volatility function, which is constructed based on stochastic collocation [2, 27, 4, 18, 22] and, as a consequence, admits a perfect calibration to arbitrage-free European-type option prices. The kernel process can be any stochastic process —in the case however that the moments of the kernel variable exist and are numerically stable, optimal collocation points can be determined by which the local volatility function is defined [11].

The CLV model allows for flexibility regarding the forward smile. The forward smile is governed by the kernel process and reflects the transition densities between future states of the underlying, which determine the price of a path-dependent product [3]. By an appropriate choice of the kernel process and its parameter values, the CLV model is well-capable of pricing exotic options, while maintaining a fit to liquid European-style options. In this article we consider three different kernel processes: the Ornstein-Uhlenbeck (OU) and Cox-Ingersoll-Ross (CIR) processes and the Heston model.

Another advantageous property of the CLV model is the fact that the local volatility function only needs to be evaluated at the time-points of interest. In addition, the kernel process typically allows for

*Corresponding author
large time-steps in a simulation. This particularly holds if an analytical solution is available (as for e.g. the OU and CIR kernel processes), however also for other processes efficient simulation schemes exist, e.g. the Heston model can be efficiently simulated by Andersen’s QE scheme [1].

The CLV model, by its flexibility in controlling the forward smile and its rapid Monte Carlo evaluation, allows for an efficient Monte Carlo calibration to exotic options, while the fit to European-type options is preserved.

2. The Collocating Local Volatility Model. In this section we discuss the main characteristics of the Collocating Local Volatility (CLV) model [12]. Also, we write the model in a standard form and derive its pricing PDE.

The CLV model is represented as follows, under the risk-neutral $\mathbb{Q}$-measure:

\begin{align}
S(t) &= g_N(t, X(t)), \\
\text{d}X(t) &= \frac{\text{d}\mathbb{Q}(t, X(t))}{\mathbb{Q}(t, X(t))} + \sigma(t, X(t)) \text{d}W^Q(t), \quad X(t_0) = X_0,
\end{align}

where $g_N(\cdot, X(\cdot)) : [t_0, T] \times \mathbb{R}^n \rightarrow \mathbb{R}$, $\mathbb{P}(\cdot, X(\cdot)) : [t_0, T] \times \mathbb{R}^n \rightarrow \mathbb{R}$, $\mathbb{Q}(\cdot, X(\cdot)) : [t_0, T] \times \mathbb{R}^n \rightarrow \mathbb{R}^{n \times d}$ and $W^Q(\cdot)$ is an $n$-dimensional Brownian motion. The model consists of two elements that are evaluated separately. The first building block is the kernel process $X(\cdot)$ in (2). The second building block is the CLV element (1), which connects the kernel process to liquid market quotes via the local volatility function $g_N(\cdot, \cdot)$, which is based on the stochastic collocation method [2, 27, 4, 18, 22].

An advantageous property of the CLV model is that, by construction, function $g_N(\cdot, \cdot)$ guarantees an almost perfect calibration to arbitrage-free European-type option prices, independently of the kernel parameter values. Basically this function, given liquid market quotes for expiries $T_1, \ldots, T_M$, yields a highly accurate interpolation through the pairs $(x_{i,j}, s_{i,j})$, $i = 1, \ldots, M$, $j = 1, \ldots, N$, with $x_{i,1}, \ldots, x_{i,M}$ and $s_{i,1}, \ldots, s_{i,N}$ representing the collocation points and collocation values corresponding to $T_i$, respectively. The collocation points may be established based on the moments of the kernel variable. The collocation values are computed by

\[ s_{i,j} = F_{1 - S(T_i)}^{-1}(F_X(T_i)(x_{i,j})) , \]

where the cumulative distribution function of $S(T_i)$ under the risk-neutral measure.

The function $g_N(\cdot, \cdot)$ is an interpolation through the $s_{i,j}$-values, given particular $t$ and $X(\cdot)$ values. Choosing $g_N(\cdot, \cdot)$ in the Lagrange form is well-accepted in the field of Uncertainty Quantification, see e.g. [21]. However, this choice does not guarantee monotonicity in the strike direction, which is a desirable property. We therefore choose a piecewise cubic Hermite interpolation, which is guaranteed to be monotonic and continuously differentiable, see e.g. [9].

Besides for its almost perfect calibration, a second beneficial property of the CLV model is the fact that, in e.g. a Monte Carlo simulation framework, we do not need to evaluate $g_N(\cdot, \cdot)$ at each time-step, which is the case for the standard Local Volatility model [8, 7]. For example, for pricing a European-type option we simulate the kernel process (2) up to the option’s maturity $T$ and subsequently compute $g_N(T, X(T))$ (1). In the case that the time-points of interest are specified on a coarse grid, we prefer a simulation method for the kernel variable which is low-biased for large time-steps. Moreover, in the case that the kernel process has an analytical solution (e.g. the Ornstein-Uhlenbeck and Cox-Ingersoll-Ross processes), it allows for an exact simulation method with large time-steps.

A third advantageous characteristic of the CLV model is its flexibility in controlling the forward smile, while maintaining an almost perfect fit to European-type options by construction. It is a well-known property of the standard Local Volatility model that it produces a flattening forward smile, which may be not in line with market observations [24]. This may lead to a mispricing of products that are sensitive to the forward smile, like cliquets and barrier options. Regarding the CLV model, however, as discussed in [12], the autocorrelation of the kernel process affects the forward smile. As such, the choice of kernel process and the kernel parameter values determine the forward smile generated by the CLV model, without affecting the almost perfect calibration to European-type options. The payoff of a path-dependent product is determined by the evolution of the underlying through time, i.e. its price depends on the transition densities from one future state to another [3].
As the CLV model is flexible in controlling the forward smile and can be efficiently evaluated, it allows for an efficient Monte Carlo calibration to exotic options, while the fit to European-type options is maintained. We consider the forward smile for three different choices of the kernel process, moreover we calibrate the kernel process to FX barrier options.

References


Higher order approximation of call option prices under stochastic volatility models

ARCHIL GULISASHVILI
DEPARTMENT OF MATHEMATICS, OHIO UNIVERSITY, ATHENS OH 45701.
EMAIL: gulisash@ohio.edu

RAÚL MERINO
FACULTAT DE MATEMÀTIQUES, UNIVERSITAT DE BARCELONA, GRAN VIA 585, 08007 BARCELONA, SPAIN
EMAIL: raul.merino85@gmail.com

MARC LAGUNAS
DEPARTMENT OF MATHEMATICS, UNIVERSITY OF OSLO, P.O. BOX 1053 BLINDERN, 0316 OSLO, NORWAY
EMAIL: marclagu@math.uio.no

JOSEP VIVES
FACULTAT DE MATEMÀTIQUES, UNIVERSITAT DE BARCELONA, GRAN VIA 585, 08007 BARCELONA, SPAIN
EMAIL: josep.vives@ub.edu

Abstract. In this talk, we find a Taylor type expansion of the price of a plain vanilla call under the Heston model. We derive closed form approximation formulas with precise error estimates of different orders that depend on the number of terms of the expansion we consider. The methodology is based on Itô calculus. With five terms we have a theoretical error of $O(\nu^3(|\rho|+\nu))$ where $\nu$ is the volatility of the volatility and $\rho$ is the correlation between the underlying price and the stochastic volatility. With twelve terms we obtain a theoretical error of $O(\nu^4)$. In particular, for the zero correlation case, we have an approximation of order $O(\nu^6)$ with only four terms. We compare the performance of our approximated formulas with previous less accurate formulas and with other methods existing in the literature, and show they are very efficient, specially for low values of the vol-vol parameter or when time to maturity is small.

1. Introduction. Stochastic volatility models were introduced in order to account for various downsides in the constant volatility assumption, on which the celebrated Black-Scholes model is based. One of the most popular stochastic volatility models is the Heston model. In [1], an approximation formula with a general error term was obtained for the call price in the Heston model. This error term was quantified in [2] and it was shown to be of order $O(\nu^2(|\rho|+\nu^2))$. In the previous expression, $\nu$ is the vol-vol parameter and $\rho$ is the correlation between the spot price and the volatility. However, in the above-mentioned approximation formula, some terms of the order $\nu^2$ were ignored, whereas other terms of the same order were kept. This may be considered as a drawback in the approximation formula obtained in [2].

In the present talk we use recursively a general decomposition of a price obtained in [4] to obtain different approximations of the call price. With five terms we have a theoretical error of $O(\nu^3(|\rho|+\nu))$ where $\nu$ is the volatility of the volatility and $\rho$ is the correlation between the underlying price and the stochastic volatility. With twelve terms we obtain a theoretical error of $O(\nu^4)$. In particular, for the zero correlation case, we have an approximation of order $O(\nu^6)$ with only four terms.

2. Preliminaries and notations. Let $S = \{S_t, t \in [0,T]\}$ be a strictly positive price process under a market chosen risk neutral probability. We consider the change of $X_t = \log S_t$, $t \in [0,T]$ that follows the process

$$dX_t = \left(r - \frac{1}{2} \sigma_t^2\right) dt + \sigma_t \left(p dW_t + \sqrt{1-p^2} d\tilde{W}_t\right),$$

with initial condition $x_0 = \log s_0$. 

165
The following notation is used in the talk:

- The Black-Scholes function is denoted by \((BS)\).
- The call option price in the Black-Scholes model is given by \(V_t = e^{-r(T-t)}E_t[(e^{X_T} - K)^+]\).
- We define \(M_t = \int_0^T E_t[\sigma_u^2] \, ds\).
- We define \(v_t^2 := E_t(\sigma_t^2) = \frac{1}{T-t} \int_t^T E_t[\sigma_u^2] \, ds\).
- We define the operators \(\Lambda := x \partial_x\), \(\Gamma := x^2 \partial_x^2\) and \(\Gamma^2 = \Gamma \circ \Gamma\).
- Given two continuous semimartingales \(X\) and \(Y\), we define \(L[X,Y]_t := E_t \left[ \int_t^T \sigma_u d[X,Y]_u \right]\) and \(D[X,Y]_t := E_t \left[ \int_t^T d[X,Y]_u \right]\).

3. General Expansion Formulas. In [1], the following decomposition is obtained:

**Theorem 3.1. (BS expansion formula)** For every \(t \in [0,T]\),

\[
V_t = (BS)(t, X_t, v_t) + \frac{\rho}{2} E_t \left[ \int_t^T e^{-r(u-t)} \Lambda \Gamma (BS)(u, X_u, v_u) \sigma_u dW[M]_u \right] + \frac{1}{8} E_t \left[ \int_t^T e^{-r(u-t)} \Gamma^2 (BS)(u, X_u, v_u) d[M,M]_u \right] 
\]

This formula is a particular case of the following result that comes from [4].

**Theorem 3.2.** Suppose the functional \(A_t := A(t, X_t, v_t^2), t \in [0,T]\), satisfies the Black-Scholes equation, and \(B\) is adapted to the filtration \(\mathcal{F}^W\). Then, for every \(t \in [0,T]\),

\[
e^{-r(T-t)}E_t[A_T B_T] = A_t B_t + \frac{\rho}{2} E_t \left[ \int_t^T e^{-r(u-t)} \Lambda \Gamma A_u B_u \sigma_u dW[M]_u \right] + \frac{1}{8} E_t \left[ \int_t^T e^{-r(u-t)} \Gamma^2 A_u B_u d[M,M]_u \right] + \rho E_t \left[ \int_t^T e^{-r(u-t)} \Lambda A_u \sigma_u dW[B]_u \right] + \frac{1}{2} E_t \left[ \int_t^T e^{-r(u-t)} \Gamma A_u d[M,B]_u \right] + E_t \left[ \int_t^T e^{-r(u-t)} A_u dB_u \right].
\]

Unfortunately, the terms of Theorem 3.1. can be cumbersome to compute. We search for an alternative approximative formula with terms easier to compute. The starting point in the construction of an infinite expansion of \(V_t\) is the formula in (2). In [1], Theorem 3.2. is applied to the formula in (2), finding two main terms and discarding the others. The main idea is that instead of discarding terms, Theorem 3.2. can be iteratively used for each one of them. Performing this process to each new term, a series of infinite terms is obtained. By selecting accurately the terms the theoretical error of the approximation can be controlled.

**Remark 3.3.** In [3], an exact representation of a conditional expectation is given in terms of forest of iterated integrals, also called diamonds. The expansion of the call option price found in the present section coincides with the obtained in [3].
4. Price approximation formulas under the Heston Model. The log-price process \( X \) in the Heston model satisfies the following system of stochastic differential equations:

\[
\begin{align*}
\frac{dX_t}{t} &= \left( r - \frac{\sigma_t^2}{2} \right) dt + \sigma_t \left( \rho dW_t + \sqrt{1-\rho^2} d\tilde{W}_t \right), \\
\frac{d\sigma_t^2}{t} &= \kappa (\theta - \sigma_t^2) dt + \nu \sqrt{\sigma_t^2} dW_t.
\end{align*}
\]

We start with an assertion that provides an approximation of the order \( O(\nu^3(|\rho| + \nu)) \).

**Theorem 4.1. (2nd order approximation formula)** For every \( t \in [0, T] \),

\[
V_t = (BS)(t, X_t, v_t) + \Gamma^2(BS)(t, X_t, v_t) \left( \frac{1}{8} D[M, M]_t \right) \\
+ \Lambda \Gamma(BS)(t, X_t, v_t) \left( \frac{\rho}{2} L[W, M]_t \right)^2 \\
+ \rho \Lambda^2 \Gamma(BS)(t, X_t, v_t) L[W, \frac{\rho}{2} L[W, M]_t] + \epsilon_t,
\]

where \( \epsilon_t \) is the error term satisfying

\[
|\epsilon_t| \leq \nu^3 (|\rho| + |\rho|^3 + \nu) \left( \frac{1}{T} \wedge (T - t) \right) \Pi(\kappa, \theta),
\]

where \( \Pi(\kappa, \theta) \) a positive function.

The next assertion contains an approximation formula with the error term of the form \( O(\nu^4) \).

**Theorem 4.2. (3rd order approximation formula)** For every \( t \in [0, T] \),

\[
V_t = (BS)_t + \Lambda \Gamma(BS)_t \left( \frac{\rho}{2} L[W, M]_t \right)^2 \\
+ \frac{1}{6} \Lambda^2 \Gamma^3(BS)_t \left( \frac{\rho}{2} L[W, M]_t \right)^3 + \Lambda \Gamma^3(BS)_t \left( \frac{\rho}{2} L[W, M]_t \right) \left( \frac{1}{8} D[M, M]_t \right) \\
+ \rho \Lambda^2 \Gamma(BS)_t L[W, \frac{\rho}{2} L[W, M]_t] + \rho \Lambda \Gamma^2(BS)_t L[W, \frac{1}{8} D[M, M]_t] \\
+ \frac{1}{2} \Lambda^3 \Gamma(BS)_t D[M, \frac{1}{8} D[M, M]_t] + \frac{1}{2} \Lambda \Gamma^2(BS)_t D[M, \frac{\rho}{2} L[W, M]_t] \\
+ \rho \Lambda \Gamma^3(BS)_t \left( \frac{\rho}{2} L[W, M]_t \right)^2 + \rho \Lambda^2 \Gamma(BS)_t L[W, \rho L[W, \frac{\rho}{2} L[W, M]_t] \\
+ \Gamma^2(BS)_t \left( \frac{1}{8} D[M, M]_t \right) + \epsilon_t.
\]

where the error term \( \epsilon_t \) is at least \( O(\nu^4) \).

In the case, where the Heston model is uncorrelated, we can obtain a similar expansion with fewer terms and a better error estimate.

**Corollary 4.3.** If \( \rho = 0 \), then

\[
V_t = (BS)_t + \Gamma^2(BS)_t \left( \frac{1}{8} D[M, M]_t \right)^2 + \frac{1}{2} \Gamma^3(BS)_t D[M, \frac{1}{8} D[M, M]_t] + \epsilon_t
\]

where \( \epsilon_t \) is the error term satisfying

\[
|\epsilon_t| \leq \nu^6 \left( \frac{1}{T} \wedge (T - t) \right) \Pi(\kappa, \theta),
\]

where \( \Pi(\kappa, \theta) \) a positive function.

**Remark 4.4.** Call price approximations similar to those formulated above for the Heston model can also be obtained for the Bates model. See [4].
Higher order approximation of call option prices under stochastic volatility models

5. Numerical results. We compare the performance of the call price approximation formula proposed in [1] and [2] with the new approximation formulas presented in this talk. As a benchmark price, we choose a Fourier transform based pricing formula. In particular, we use a semi-closed form solution with one numerical integration as a reference price. As an example, in Figure 1, we deal with the case, where \( \nu \) is small, while \(|\rho|\) is close to one. We observe that the new approximation formulas perform clearly better than the previously known one. The approximation error is in the range \( 10^{-4} - 10^{-8} \) for the formula in which the error term is of order \( O(\nu^2) \), and for the formula with the error term of order \( O(\nu^3) \), the approximation error is in the range \( 10^{-7} - 10^{-10} \).

One of the main advantages of the proposed pricing approximations is their computational efficiency. We observe that the approximations of order \( O(\nu^2) \) and \( O(\nu^3) \) are around 43-45 times faster than the approximation based on the fast Fourier transform methodology, while the approximation of order \( O(\nu^4) \) is around 36 times faster than the latter one. Therefore, the approximations of order \( O(\nu^2) \) and \( O(\nu^3) \) are around 1.14-1.25 times less time-consuming than the \( O(\nu^4) \)-approximation.

References
Pricing and Hedging of Guaranteed Minimum Benefits using Power Series Approximation Techniques

Nikolay Gudkov
School of Risk and Actuarial Studies, UNSW Sydney, Australia
Email: n.gudkov@unsw.edu.au

Jonathan Ziveyi
School of Risk and Actuarial Studies, UNSW Sydney, Australia
Email: j.ziveyi@unsw.edu.au

Abstract. Majority of the developed economies are experiencing population ageing with people living longer due to advances in medicines and lifestyle quality. Such developments have been putting a lot of pressure on governments and pension fund providers who are exposed to the resulting longevity risk. Variable annuities (VAs) constitute a class of financial products designed to tackle challenges associated with both investment and longevity risk. These contracts enable policyholders to participate in financial markets via linked funds and at the same time provide protection against long-term life contingencies. In this paper, we devise a numerical technique for pricing Guaranteed Minimum Benefit (GMB) riders embedded in variable annuities. The method utilises multidimensional transforms of the characteristic function for the underlying stochastic process and enables to express solutions of pricing partial differential equations in terms of power series with coefficients known in the closed form. Our results demonstrate the high computational efficiency of the series approximation method for the computation of prices and hedge ratios of GMBs under the stochastic volatility and stochastic interest rate modelling framework. The findings of the paper can help insurers with efficient quantification of various risks associated with GMBs in variable annuities.

1. Introduction. Pricing of European option involve calculation of the discounted expected value of the payoff of the contract at maturity under the risk-neutral probability such that

\[ V_t = E_Q^t \left[ e^{-\int_t^T r_s ds} V_T \right]. \]

In this paper, we aim to extend the results presented in [3] who proposed to use a power series approximation in order to address this problem. The authors derive an approximate formula to the price of a European option under the Heston stochastic volatility and the Cox-Ingersoll-Ross stochastic interest rate modelling framework herewith abbreviated HCIR, when maturity of the contract, \( T \leq 1 \). We develop a power series solution which enables us to find the value of the characteristic function of the HCIR process and perform pricing of European option with maturities, \( T > 0 \). We demonstrate superior accuracy and computational efficiency of the method in comparison to other existing techniques (e.g. Monte-Carlo simulations) when applying it to valuation of European options and Guaranteed Minimum Benefits (GMABs) in Variable Annuities.

2. Methodology.

2.1. Valuation of European call option. Suppose that under the risk-neutral measure \( Q \) the dynamics of the underlying asset price \( S_t \) follows the HCIR model

\[ dS_t = r_t S_t dt + \sqrt{\nu_t} S_t dW^1_t, \quad d\nu_t = \kappa(\theta - \nu_t) dt + \sigma \sqrt{\nu_t} dW^2_t, \quad dr_t = \alpha(\beta - r_t) dt + \eta \sqrt{r_t} dW^3_t, \]

where \( \nu_t \) is variance, \( r_t \) is risk-free interest rate; \( \kappa \) is a speed of mean-reversion and \( \theta \) is a long-term mean of the stochastic variance process, while \( \alpha \) is a speed of mean-reversion and \( \beta \) is a long-term mean of the stochastic interest rate process. We assume that the \( W^1_t, W^2_t \) and \( W^3_t \) are standard Brownian motions with the following correlation structure: \( dW^1_t dW^2_t = \rho dt, \quad dW^1_t dW^3_t = dW^2_t dW^3_t = 0 \).
Pricing and Hedging of Guaranteed Minimum Benefits

Under the forward measure, $Q^T$, where the zero-coupon Bond, $P(t,T,r)$, is used as a numeraire, the price of a European option can be expressed as

$$U(y,v,r) = P(r,t,T)(P_1 - KP_2),$$

(2)

where

$$P_1 = f(-i; t,T,y,v,r) \left( \frac{1}{2} + \frac{1}{\pi} \int_0^\infty \text{Re} \left[ \frac{e^{-i\phi\ln(K)} f(\phi - i; t,T,y,v,r)}{i\phi f(-i; t,T,y,v,r)} \right] d\phi \right),$$

(3)

and

$$P_2 = \frac{1}{2} + \frac{1}{\pi} \int_0^\infty \text{Re} \left[ \frac{e^{-i\phi\ln(K)} f(\phi; t,T,y,v,r)}{i\phi} \right] d\phi.$$  

(4)

The characteristic function $f(\phi; t,T,y,v,r)$ of the process in Eq (1) is assumed to have an exponentially affine structure:

$$f(\phi; \tau,y,v,r) = e^{C(\phi,\tau) + D(\phi,\tau)v + E(\phi,\tau)r + i\phi y},$$

(5)

with $\tau = T - t$. Here, functions $C(\phi,\tau)$, $D(\phi,\tau)$ and $E(\phi,\tau)$ should solve the following system of ODEs:

$$\frac{dD(\phi,\tau)}{d\tau} = \frac{1}{2} \sigma^2 D(\phi,\tau)^2 + (i\phi\rho\sigma - \kappa)D(\phi,\tau) - \frac{1}{2}(i\phi + \phi^2),$$

(6)

$$\frac{dE(\phi,\tau)}{d\tau} = \frac{1}{2} \eta^2 E(\phi,\tau)^2 - (\alpha + B(\tau)\eta^2)E(\phi,\tau) + i\phi,$$

(7)

$$\frac{dC(\phi,\tau)}{d\tau} = \kappa \theta D(\phi,\tau) + \alpha \beta E(\phi,\tau)$$

(8)

subject to the initial conditions $D(\phi,0) = C(\phi,0) = E(\phi,0) = 0$. Eq (6) is a Riccati equation with constant coefficients, whose solution can be found in closed-form as

$$D(\tau) = \frac{(d + (i\phi\rho\sigma - \kappa))(1 - e^{d\tau})}{\sigma^2(1 - ge^{d\tau})},$$

(9)

where $g = \frac{(i\phi\rho\sigma - \kappa) - d}{(i\phi\rho\sigma - \kappa) + d}$. Eq (7) is a Riccati equation with non-constant coefficient and the solution to it cannot be found analytically. We perform a transformation of Eq (7) to a second order differential equation by substituting

$$E(\phi,\tau) = -\frac{2U''(\phi,\tau)}{\eta U(\phi,\tau)},$$

(10)

such that

$$U''(\phi,\tau) + (\alpha + B(\tau)\eta^2)U'(\phi,\tau) + \frac{i\phi\eta^2}{2}U(\phi,\tau) = 0.$$  

(11)

The solution to the second order ODE (11) can be represented as a power series

$$U(\phi,\tau) = \sum_{n=0}^\infty a_n^{(l)}(\tau - \tilde{\tau})^n,$$

(12)

where the point of expansion, $\tilde{\tau}$, is chosen in a way that the whole series converges. For this, we have to consider that the radius of convergence of the series in Eq (12) is as large as the distance between $\tilde{\tau}$ and the nearest singularity point of ODE (11), see [2] and [3].

We fix the number of terms in the series in Eq (12) to $N < \infty$, and solve a system of linear equations:

$$M\hat{a}^{(l)} = b,$$

(13)
where the matrix $M \in \mathbb{R}^{N \times N}$ and the vector $b \in \mathbb{R}^N$ are defined as follows:

$$M = \begin{pmatrix}
\Gamma_{(1,+)} + \Theta_{(1,+)} + \Psi_{(1,+)} + \Omega_{(1,+)} \\
\Gamma_{(2,+)} + \Theta_{(2,+)} + \Psi_{(2,+)} + \Omega_{(2,+)} \\
\vdots \\
\Gamma_{(n,+)} + \Theta_{(n,+)} + \Psi_{(n,+)} + \Omega_{(n,+)} \\
M_{(N-2,+)} \\
M_{(N-1,+)} \\
M_{(N,+)}
\end{pmatrix},
\begin{pmatrix}
-\frac{1}{2}i\phi\eta^2(\alpha + m)e^{m\bar{\tau}}c_1 \\
-\frac{1}{2}i\phi\eta^2(\alpha + m)e^{m\bar{\tau}}c_2 \\
\vdots \\
-\frac{1}{2}i\phi\eta^2(\alpha + m)e^{m\bar{\tau}}c_a \\
-\frac{1}{2}i\phi\eta^2(\alpha + m)e^{m\bar{\tau}}c_{N-2} \\
0 \\
-\frac{1}{2}i\phi\eta^2\Psi
\end{pmatrix}. $$

Here, for $1 \leq n \leq N-2$,

$$\Gamma_{(n,+)} = \begin{pmatrix}
0, 0, \ldots, 0, \frac{1}{2}i\phi\eta^2\Psi, \Lambda(n + 1), \Psi(n + 2)(n + 1), 0, \ldots, 0
\end{pmatrix},$$

$$\Theta_{(n,+)} = (\alpha + m)e^{m\bar{\tau}}(0, 2e_n, 6e_{n+1}, \ldots, k(k-1)c_{n+2-k}, \ldots, (n+1)e_n1, 0, \ldots, 0),$$

$$\Psi_{(n,+)} = \Xi(e_n, 2e_n-1, \ldots, k(k-1)e_n-k, \ldots, (n+1)e_n1, 0, \ldots, 0),$$

$$\Omega_{(n,+)} = \frac{1}{2}i\phi\eta^2(\alpha + m)e^{m\bar{\tau}}(e_{n-1}, e_{n-2}, \ldots, e_{n-k}, \ldots, c_1, 0, \ldots, 0),$$

$$M_{(N-2,+)} = (1, 2(-\bar{\tau}), \ldots, n(-\bar{\tau})^{n-1}, \ldots, N(-\bar{\tau})^{N-1}),$$

and

$$M_{(N,+)} = (\Lambda, 2\Psi, 0, 0, \ldots, 0).$$

Solution to Eq (8) for $C(\phi, \tau)$ can be found by direct integration of Eq (9) and Eq (10):

$$C(\phi, \tau) = \frac{\kappa\theta}{\sigma^2} \left( \lambda - (i\phi\sigma - \kappa)\tau - 2\ln \left( \frac{1 - ge^{\lambda\tau}}{1 - g} \right) \right) - \frac{2\alpha\beta}{\eta^2} \int_0^\tau \frac{U'(\phi, s)}{U(\phi, s)} ds. \quad (20)$$

The second term in the equation above can be written as

$$\frac{2\alpha\beta}{\eta^2} \int_0^\tau \frac{U'(\phi, s)}{U(\phi, s)} ds = \frac{2\alpha\beta}{\eta^2} \sum_{k=0}^{l-1} \left\{ \int_{\tau_k}^{\tau_{k+1}} \sum_{n=0}^{\infty} b_n^k(s - \bar{\tau})^n \right\} + \frac{2\alpha\beta}{\eta^2} \left\{ \int_{\tau_n}^{\tau_{n+1}} \sum_{n=0}^{\infty} b_n^0(s - \bar{\tau})^n \right\},$$

$$= \frac{2\alpha\beta}{\eta^2} \sum_{k=0}^{l-1} \left\{ \sum_{n=0}^{\infty} b_k^n(\tau_{k+1} - \bar{\tau})^n + \frac{2\alpha\beta}{\eta^2} \left\{ \sum_{n=0}^{\infty} b_0^n(\tau - \bar{\tau})^n \right\} \right\}, \quad (21)$$

where $b_0^n = a_0^{(l)}(l)\sum_{j=0}^{n} a_j^{(l)}$, and $b_k^n = \frac{(k+1)a_k^n}{n+1}$, for $k \geq 1$.

### 2.2. Valuation of the GMAB

Assume that under $\mathbb{Q}$ the dynamics of the investment account in the Guaranteed Minimum Accumulation Benefit, $A_t$, is

$$dA_t = \begin{cases}
A_t \frac{dS_t}{S_t} - \lambda A_t dt, & \text{if } A_t > 0 \\
0, & \text{otherwise}
\end{cases} \quad (22)$$

with $A_0 = P$ being the insurance premium and $\lambda$ being the insurance fee paid by the policyholder. We adopt the CIR-type model for the force of mortality, $\mu_t$ (see [1]):

$$d\mu_t = \xi(\bar{\mu}_t - \mu_t) dt + \sigma_\mu \sqrt{\mu_t} dW_t^\mu, \quad (23)$$

where $\bar{\mu}_t = e^{-\xi_2(t + T)^{\alpha_2}}$, and $dW_t^\mu$ is the standard Brownian motion independent of $W_t^1$, $W_t^2$ and $W_t^3$. Under this modelling setting the value of the GMAB option is given by the equation:

$$P = A_0 = E^Q_t \left[ e^{-\int_0^T \mu_s ds} \right] E^Q_t \left[ e^{\int_0^T \tau_s ds} \max \{A_T(\lambda), P \} \right], \quad (24)$$

which has to be solved subject to the insurance fee, $\lambda$, implicitly.
3. Numerical results. In our numerical experiments we use the following set of base parameters $T = 5$, $S_0 = 100$, $K = 100$, $v_0 = 0.05$, $r_0 = 0.03$, $\kappa = 10$, $\theta = 0.2$, $\sigma = 0.1$, $\rho = 0.1$, $\alpha = 2$, $\beta = 0.05$, $\eta = 0.05$. Figure (1a) demonstrates that convergence of the solution is achieved when $N \geq 15$ terms are included into the power series, Eq (12). From Figure (1b) we notice that the results provided by the power series approximation method are very close to the ones obtained via Monte Carlo Simulations (top plot), but computational efficiency of the former method is much higher (bottom plot).

(a) Absolute error in the option price as a function of number of terms, $N$, in the series in Eq (12).

(b) Comparison of the European option prices obtained via different methods (power series approximation and Monte Carlo simulations).

Figure 2: Fair insurance fee for the GMAB option. Mortality modelling parameters: $c_1 = 90.43$, $c_2 = 10.36$, $\xi_\mu = 0.5$, $\sigma_\mu = 0.03$.

References


Regression likelihood ratio method Monte Carlo sensitivities

Jonathan Hüser
RWTH Aachen University
Email: huesser@stce.rwth-aachen.de

Uwe Naumann
RWTH Aachen University
Email: naumann@stce.rwth-aachen.de

1. Introduction. The payoff function of financial derivatives can be discontinuous in the price of the underlying and payoffs often have discontinuous first derivatives. Payoff irregularity makes pathwise (second-order) differentiation in Monte Carlo estimators challenging and has motivated the development of many sophisticated methods for various different settings [1, 2, 3, 5, 9]. In this abstract we present a new Monte Carlo method for the estimation of parameter sensitivities (Greeks) that works with discontinuous payoff functions.

We consider the case of an underlying where price $S(t)$ is modeled by a univariate stochastic differential equation (SDE)

$$dS(t) = a(S(t), \theta)dt + b(S(t), \theta)dW(t)$$

for $t \in [0, T]$ with $S(0) = S_0(\theta)$, $\theta \in \mathbb{R}^d$ the parameters and $W(t)$ the standard Brownian motion. Derivative price is computed as the expected payoff $V = E[f(S(T))]$ at expiration time $T$ (for brevity we omit explicitly writing the discount factor). The payoff function $f \in L^1$ is assumed to be weakly differentiable, i.e. $f$ can have jump discontinuities but we can still interchange differentiation and integration. The presented method also extends to the multivariate case (ongoing work that is not discussed in the scope of this abstract).

The likelihood ratio method is an approach to computing Monte Carlo sensitivities based on differentiating the probability density function of a random variable [6]. Let $p(\cdot|\theta)$ be the probability density function of $S(T)$. With an iid sample $S^i \sim p(S^i|\theta)$ for $i = 1, \ldots, N$ we get the unbiased Monte Carlo estimator $\hat{V} = N^{-1} \sum_{i=1}^N f(S^i)$ with $E[\hat{V}] = V$. If $p(S(T)|\theta)$ is known explicitly the likelihood ratio method can be used to compute sensitivities as an expectation over the same distribution as follows. The likelihood ratio method uses the score function

$$\partial_\theta \log(p(S(T)|\theta)) = \frac{\partial_\theta p(S(T)|\theta)}{p(S(T)|\theta)}$$

(2)

to turn the differentiated integral into an expectation as follows

$$\partial_\theta E[f(S(T))] = \int_{-\infty}^{\infty} f(S(T)) \partial_\theta \log(p(S(T)|\theta))dS(T)$$

(3)

$$= \int_{-\infty}^{\infty} f(S(T)) \partial_\theta \log(p(S(T)|\theta))p(S(T)|\theta)dS(T)$$

(4)

$$= E[f(S(T)) \partial_\theta \log(p(S(T)|\theta))]$$

(5)

for $l = 1, \ldots, d$. With an iid sample $S^i \sim p(S^i|\theta)$ for $i = 1, \ldots, N$ we get the unbiased Monte Carlo estimator

$$\hat{\partial_\theta V} = N^{-1} \sum_{i=1}^N f(S^i) \partial_\theta \log(p(S^i|\theta))$$

(6)

with $E[\hat{\partial_\theta V}] = \partial_\theta V$.

It is, however, typical that we do not know explicitly how to evaluate $p(S^i|\theta)$ and sampling is only done approximately by simulating the stochastic differential equation. In this abstract we assume that
Monte Carlo simulation is done via the Euler-Maruyama discretization scheme with \( n \) steps

\[
S_{j+1} = S_j + a(S_j, \theta)h + b(S_j, \theta)\sqrt{h}X_j
\]

where \( h = T/n \) is the time step size and \( X_j \sim \phi(X_j) \) is a standard normal random variable for \( i = 0, \ldots, n - 1 \). The presented method also generalizes to other discretization schemes (ongoing work that is not discussed in the scope of this abstract). If a simulation algorithm is used based on an iid sample of \( X^i \sim \phi(X^i) \) we can write the final time step as a parametric function \( S_n^i = g(\theta, X^i) \) for \( i = 1, \ldots, N \). The pathwise sensitivity method works by differentiating the expectation over \( X \sim \phi(X) \) where the distribution of \( X \) does not depend on \( \theta \)

\[
\nabla_\theta E[f(g(\theta, X))] = E\left[\nabla_\theta f(g(\theta, X))\nabla_\theta g(\theta, X)\right] .
\]

If \( f \) is Lipschitz continuous we have \( \partial S_n f \in L^1 \) and the above expectation can be estimated from an iid sample \( X^i \sim \phi(X) \) for \( i = 1, \ldots, N \). In that case we can efficiently compute the whole gradient \( \nabla_\theta g(\theta, X^i) \) of each Monte Carlo path by adjoint algorithmic differentiation (AAD) [4, 7].

In this abstract we show how to combine the likelihood ratio method with pathwise AAD in a new way that allows it to work for discontinuous payoffs in a setting where \( p(S(T)|\theta) \) is not known explicitly. The presented approach is comparable to the Vibrato Monte Carlo method by Giles [5]. But in contrast to the Vibrato method the variance of our proposed Monte Carlo estimator is independent of the discretization step size \( h \) (i.e. for a constant number of simulated paths the expected error does not increase when decreasing the discretization bias).

2. Regression likelihood ratio method. In this section we present the main result in the form of a theorem. Theorem 2.1. proposes a way of rewriting \( \nabla_\theta E[f(g(\theta, X))] \) as an expectation that can be estimated via the Monte Carlo method. The theorem uses a combination of different simulation path sensitivities and the partial score function \( \partial S_n \log(\hat{p}(S_n)) \). The function \( \hat{p} \) is the probability density of the final time step value \( S_n \) of the simulation and, hence, an approximation of \( p \). The partial score function \( \partial S_n \log(\hat{p}(S_n)) \) is usually not known explicitly. We later also show how the score function can be robustly approximated by regression based on an iid sample \( S^i_n \sim \hat{p}(S^i_n) \) for \( i = 1, \ldots, N \).

Theorem 2.1.

\[
\nabla_\theta E[f(g(\theta, X))] = E\left[f(g(\theta, X))\left(-\nabla_\theta \log(\partial X_n g(\theta, X)) - \nabla_\theta g(\theta, X)\partial S_n \log(\hat{p}(g(\theta, X)))\right)\right] .
\]

Proof. The proof is omitted for space reasons.

The factor

\[
-\nabla_\theta \log(\partial X_n g(\theta, X)) - \nabla_\theta g(\theta, X)\partial S_n \log(\hat{p}(g(\theta, X)))
\]

(10)
gives an implicit way of evaluating the unknown score function \( \nabla_\theta \log(\hat{p}(g(\theta, X)|\theta)) \). So Monte Carlo estimation of the right-hand side of Equation (9) represents a likelihood ratio method for simulated \( S^i_n = g(\theta, X^i) \sim \hat{p}(S^i_n|\theta) \). Since \( \partial S_n \log(\hat{p}(S_n)) \) is obtained via regression we call the resulting method the regression likelihood ratio method even though it is essentially a hybrid between the likelihood ratio and pathwise sensitivity method. The required gradients of \( g(\theta, X) \) and \( \log(\partial X_n g(\theta, X)) \) can be efficiently evaluated by AAD.

In order to find an approximation \( R_\eta(z) \) of \( \partial_\eta \log(\hat{p}(z)) \) we use regression as suggested by Sasaki et al. [8]. The objective is to minimize the mean squared error

\[
\int_{-\infty}^{\infty} (R_\eta(z) - \partial_\eta \log(\hat{p}(z)))^2 \hat{p}(z)dz \to \min .
\]

(11)

Multiplying out, dropping the part that is constant in \( \eta \), rewriting the score function and then via integration by parts we get

\[
\int_{-\infty}^{\infty} R_\eta(z)^2 \hat{p}(z)dz - 2\left[R_\eta(z)\hat{p}(z)\right]_{-\infty}^{\infty} + 2\int_{-\infty}^{\infty} \partial_\eta R_\eta(z)\hat{p}(z)dz \to \min
\]

(12)
where the second term is zero because \( \tilde{p} \) is the distribution of a Brownian motion based SDE simulation for time \( T < \infty \). The other two terms are expectations that can be estimated based on an iid sample drawn from \( \tilde{p} \).

The \( l^2 \)-norm regularized optimization problem resulting from the Monte Carlo estimator based on the iid sample \( z^i \) for \( i = 1, \ldots, N \) is

\[
\min_\eta \mathcal{L}(\eta) := \frac{1}{N} \sum_{i=1}^{N} \left( \mathcal{R}_\eta(z^i)^2 + 2\partial_z \mathcal{R}_\eta(z^i) \right) + \lambda ||\eta||_2^2 .
\]

(13)

To allow for efficient direct estimation we choose a regression function that is linear in the parameter \( \eta \), e.g.

\[
\mathcal{R}_\eta(z) = \sum_{i=1}^{m} \eta_i \varphi_i(z), \quad \partial_z \mathcal{R}_\eta(z) = \sum_{i=1}^{m} \eta_i \partial_z \varphi_i(z)
\]

(14)

where \( \varphi_i(\cdot) \) are basis functions equally spaced on a grid. Then the above optimization problem is quadratic and can be solved via Newton’s method for the cost of a single \( m \times m \) linear system solve.

3. Case Study. In this section we present promising first numerical results for the regression likelihood ratio method. We consider a digital option with discontinuous payoff on a single underlying asset. For this case study the Greeks are known analytically and can be used for validation.

The spot price \( S(t) \) of the underlying follows a geometric Brownian motion

\[
dS(t) = rS(t)dt + \sigma S(t)dW(t)
\]

(15)

for \( t \in [0, T] \) with \( S(0) = S_0 \), \( r \) is the risk-free rate, \( \sigma \) the volatility and \( W(t) \) the standard Brownian motion. The probability density function of \( S(T) \) at expiration time \( T \) is a log-normal distribution

\[
p(S(T)) = \frac{1}{\sqrt{2\pi} \sigma \sqrt{T}S(T)} \exp \left( -\frac{-T(r - \sigma^2/2) - \log(S_0) + \log(S(T))^2}{2\sigma^2 T} \right) .
\]

(16)

The present value of a digital option with strike \( K \) is given by the discounted expected payoff

\[
V = \exp(-rT) \int_{-\infty}^{\infty} 1_{S(T) = K} p(S(T))dS(T) = \exp(-rT) \int_{K}^{\infty} p(S(T))dS(T) .
\]

(17)

The integral can be solved analytically and the corresponding Greeks Delta \( \Delta = \partial_X V \), Rho \( \rho = \partial_r V \) and Vega \( \nu = \partial^2 V \) can be computed by symbolic differentiation.

The Euler-Maruyama discretization of a geometric Brownian motion is

\[
S_{j+1} = S_j + rS_j h + \sigma S_j \sqrt{h} X_j
\]

(18)

where \( h \) is the time step size and \( X_j \) is a standard normal random variable for \( j = 0, \ldots, n-1 \). According to Theorem 2.1. the estimators for the Greeks are given as

\[
\hat{\Delta} = \exp(-rT) \frac{1}{N} \sum_{i=1}^{N} 1_{[S_{n-1}^i - K]} \left( -\partial_r S^i_{n-1}/S^i_{n-1} + \partial_X S^i_n \mathcal{R}_\eta(S^i_n) \right)
\]

(19)

\[
\hat{\rho} = \exp(-rT) \frac{1}{N} \sum_{i=1}^{N} 1_{[S_{n-1}^i - K]} \left( -\partial_r S^i_{n-1}/S^i_{n-1} + \partial_X S^i_n \mathcal{R}_\eta(S^i_n) \right)
\]

(20)

\[
\hat{\nu} = \exp(-rT) \frac{1}{N} \sum_{i=1}^{N} 1_{[S_{n-1}^i - K]} \left( -\partial_r S^i_{n-1}/S^i_{n-1} - 1/\sigma - \partial_X S^i_n \mathcal{R}_\eta(S^i_n) \right)
\]

(21)

where \( \mathcal{R}_\eta(\cdot) \) is the approximation of \( \partial_X \log(p(\cdot)) \) obtained by regression, in this case using piecewise linear basis functions \( \varphi_i \). The sensitivities \( \partial_X S^i_n, \partial_r S^i_n, \partial_\sigma S^i_n, \partial_X S^i_{n-1}, \partial_r S^i_{n-1}, \partial_\sigma S^i_{n-1} \) are all evaluated in a single pathwise adjoint evaluation.
The above table shows the results of a numerical evaluation with three different regression grid sizes $m = 5, 10, 20$, $N = 10^4$ paths, $n = 100$ time steps, regression regularization parameter $\lambda = 10^{-4}$, expiration time $T = 2$, strike $K = 110$, risk-free rate $r = 0.1$ and volatility $\sigma = 0.2$. The regression grid minimum and maximum are chosen to span the full $S_n$ sample range. The same random seed is used for each run, i.e. the same Monte Carlo paths are used for each computation.

The estimator column shows the values obtained from using the regression likelihood ratio method. The two columns to the right show its standard deviation based on the sample statistic and its absolute error computed as the absolute value of the difference between the estimator and the analytic solution. The semi-analytic column shows the estimator result where instead of the regression $R_\eta(\cdot)$ we have used the analytic partial score function $\partial S_n \log(p(\cdot))$ and is, hence, independent of $m$. Note that the analytic partial score function is different from $\partial S_n \log(\tilde{p}(\cdot))$ that would give the unbiased-for-the-simulation result. The two columns to the right show its standard deviation and absolute error computed in the same way as described above.

We highlight the results for the regression grid size of $m = 10$ where the absolute error of all Monte Carlo sensitivity estimators are an order of magnitude smaller than their standard deviations. This comparison with the standard deviation indicates that the error made due to the regression of the partial score function is negligible compared to the sampling noise. We also see that the regression likelihood ratio method has comparable sample efficiency to the semi-analytic approach.

The choice of $m = 10$ and $\lambda = 10^{-4}$ is motivated by cross validation with a hold out set of $2 \cdot 10^3$ paths. The cross validation loss for the regression problem for different values of $\lambda$ and $m$ is listed in the following table.

<table>
<thead>
<tr>
<th>Grid size</th>
<th>$\lambda = 10^{-5}$</th>
<th>$\lambda = 10^{-4}$</th>
<th>$\lambda = 10^{-3}$</th>
<th>$\lambda = 10^{-2}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>-</td>
<td>-0.00103575</td>
<td>-0.0010385</td>
<td>-0.0010167</td>
</tr>
<tr>
<td>10</td>
<td>-0.00112392</td>
<td>-0.00112533</td>
<td>-0.00110599</td>
<td>-</td>
</tr>
<tr>
<td>20</td>
<td>-0.000779616</td>
<td>-0.000997841</td>
<td>-0.000974113</td>
<td>-</td>
</tr>
</tbody>
</table>

For a higher number of paths a higher resolution regression grid will likely be better.

4. Conclusion. In this abstract we introduced the regression likelihood ratio method for estimating Monte Carlo sensitivities for financial derivatives with discontinuous payoff functions. According to Theorem 2.1, the regression likelihood ratio method combines pathwise adjoint sensitivities [4] with the regression of a partial score function to efficiently compute likelihood ratio method sensitivities. We also demonstrated how to formulate the regularized regression problem based on a simulated Monte Carlo sample. Finally, we presented a case study where we estimate the Greeks of a digital option on an underlying that is modeled by geometric Brownian motion. The obtained numerical results are promising because they show that the method is competitive in sample efficiency with a semi-analytic likelihood ratio method.

Future work will include a comparison to the sensitivities obtained by the Vibrate Monte Carlo method [5] for varying discretization step size $h$. A more conclusive article will also contain the proof of Theorem 2.1. and how to deal with the case of multivariate stochastic differential equations and discretization methods other than the Euler-Maruyama scheme.
References


Abstract. In this paper we extend the results in Tang and Chen [4], Yu [5] and Bao et al [1] in two directions. First we find explicit closed form solutions of the pseudo MLEs using the method of Nowman [3] and we provide their asymptotic theory in the context of the CKLS [2] model characterized by a general constant elasticity of variance parameter \( \rho \). Second we obtain bias expansions for those pseudo MLEs also in terms of \( \rho \geq 0 \). We provide a general framework since only the cases with \( \rho = 0 \) and \( \rho = 0.5 \) have been considered in the literature so far. When the time series is not positive almost surely, we need to impose the restriction that \( \rho \) is a non-negative integer.

References


*The second author is very grateful for the financial support from the Spanish Ministry of Science and Innovation, project ECO2015-63845-P
Climate games: Who’s on first? What’s on second?

MARGARET INSLEY*
DEPARTMENT OF ECONOMICS, UNIVERSITY OF WATERLOO, WATERLOO, ONTARIO, CANADA
EMAIL: margaret.insley@uwaterloo.ca

PETER A. FORSYTH
CHERITON SCHOOL OF COMPUTER SCIENCE, UNIVERSITY OF WATERLOO, WATERLOO, ONTARIO, CANADA
EMAIL: peter.forsyth@uwaterloo.ca

Abstract. Strategic actions by decision makers are crucial to determining the global response to climate change. To gain insight into these strategic interactions, we contrast the standard Stackelberg game with three other possible game structures and compare with the outcome of choices by a Social Planner. In a dynamic setting, two players choose levels of carbon emissions. Rising atmospheric carbon stocks increase average global temperature which damages player utilities. Temperature is modelled as a stochastic differential equation. In addition to the Stackelberg game, we examine a game where both players act as leaders (a Leader-Leader or Trumpian game), as well as an Interleaved game where there is a significant time interval between player decisions. A solution always exists for the Stackelberg game (and the other games as well), but these may not be Nash equilibria. For the fourth game, a Nash equilibrium is chosen if it exists, and otherwise a Stackelberg game is played. One or both players may be better off in these alternative games compared to the Stackelberg game, depending on state variables. We conclude that it is important to consider alternate game structures in examining strategic interactions in pollution games. We also demonstrate that the Stackelberg game is the limit of the Interleaved game as the time between decisions goes to zero.

*The authors gratefully acknowledge funding from the Global Risk Institute, globalriskinstitute.org. ‘Who’s on first? What’s on second?’ is a reference to the famous comedy routine by Americans Bud Abbott and Lou Costello
Abstract. For several years, the Bank of Japan has implemented the quantitative easing policy (QE). Against this backdrop, Japanese short- and medium-term interest rates have stayed negative. In this study, we model the short rate as the sum of the positive part and the lower bound in order to capture negative yields. We model the former part as the quadratic form of the Gaussian process. For the latter part, we model it as the Brownian bridge on the random interval pinning at 0 on the end date of the QE where the Brownian bridge on random intervals was introduced by Bedini et al (2017). For this setting, we derive the zero coupon bond price representation by imposing the no-arbitrage condition. Furthermore, we conduct an empirical analysis based on this proposed model. Using Japanese yield curve data, we calibrate model parameters and estimate an implied exit time from the Bank of Japan’s quantitative easing.

1. Introduction. In Germany and Japan, short-term and medium-term interest rates have taken negative values against the backdrop of their quantitative easing policies. European Central Bank stopped the asset purchase program in December 2018. In Japan, although full-scale debate about reduction of quantitative easing policy has not been seen yet, the Bank of Japan keeps slowing the pace of its bond purchases. To my knowledge, there are few studies of a term structure model explicitly incorporating the exit time from the QE policy. However there are two studies to be referred. One is Marumo et al [3], the other is Futami [2]. Their models deal with the exit time from the zero interest rate policy. In Marumo et al [3], the short rate is assumed to stay at zero until the exit time arrives and it moves based on the Vasicek model after the exit time. Futami [2] deals with the general setting of Marumo et al [3]. In this work, we will construct a term structure interest rate model which captures negative interest rates and incorporates explicitly the exit time from the QE policy.

2. Setup. Let a filtered probability space \((\Omega, \mathcal{F}, (\mathcal{F}_t)_{0 \leq t}, \mathbb{P})\). The filtration \((\mathcal{F}_t)_{0 \leq t}\) satisfies the usual conditions of right-continuity and completeness. \(\mathbb{P}\) denotes the physical measure. We assume that the market is complete and has no arbitrage opportunities, so that the risk-neutral measure \(\mathbb{Q}\) uniquely exists. We deal only with a theory on \(\mathbb{Q}\) in this study. \(W_{1,2}^Q\) are independent standard Brownian motions under \(\mathbb{Q}\).

The state variable \(X_t\) satisfies the following stochastic differential equation:

\[
dX_t = K_X \left( \theta - X_t \right) dt + \Sigma_X dW_{t,1}^Q.
\]

The risk-free short rate \(r_t\) is assumed to be a quadratic form of the state variable \(X_t\) as follows:

\[
r_t = X_t' \Psi X_t + y^r_t,
\]

where \(X_t'\) represents the transposition of \(X_t\) and \(\Psi\) is assumed to be positive definite. Therefore, \(y^r_t\) can be interpreted as the lower bound of \(r_t\).

We model \(y^r_t\) as the Brownian bridge process with \(y^r_0 = 0\) and \(y^r_\tau = 0\) for \(t \geq \tau\) defined as

\[
y^r_t(\omega) = \sigma_y W_{1,2}^Q(\omega) - \frac{\sigma_y}{\tau} W_{\tau,1,2}^Q(\omega).
\]

This work was supported by JSPS Grant-in-Aid for Scientific Research (C) Grant Number 17K03802 and a Grant-in-Aid from Zengin Foundation for Studies on Economics and Finance.
where we assume that $\tau$ is a strictly positive constant value. Representing this in the SDE form,

$$
dy^\tau_t = \begin{cases} 
-\gamma^\tau_t dt + \sigma_y dW^Q_{t,2} & (t < \tau) \\
0 & (t \geq \tau)
\end{cases}
$$

$\tau$ is interpreted as the exit time from the QE policy.

### 3. Bond pricing in the case where $\tau$ is a strictly positive constant

In this section, we derive a bond pricing formula in the case where $\tau$ is a strictly positive constant. Afterwards, we denote an abnormal policy period $t < \tau$ (QE period) by a superscript of letter $a$, and a normal period, $t \geq \tau$ (post QE period) by a superscript of letter $n$. $P^n_{t,u}$ denotes the zero coupon bond price at time $t$ during the post QE period with the $u$-year maturity. $P^a_{t,u}$ denotes the zero coupon bond price at time $t$ during the QE period with the $u$-year maturity.

#### 3.1. Bond pricing in the normal period, the post-QE

In this subsection, we derive a bond pricing formula in the case where $t < \tau$. This period corresponds to the post Quantitative easing policy period.

Infinitesimal generator of $X_t$ for $t < \tau$ is provided as

$$
\mathcal{D}_t^a = (K^Q X_t (\theta^Q - X_t))' \frac{\partial}{\partial X_t} + \frac{1}{2} \text{Tr} \left( \Sigma_X \Sigma' X_t \frac{\partial^2}{\partial X_t \partial X_t'} \right).
$$

Applying the Feynman-Kac theorem to the zero coupon bond price $P^n_{t,u}$ with the maturity $T$, and $u = T - t$ leads to the following equation:

$$
\left[ \frac{\partial}{\partial t} + \mathcal{D}_t^n \right] P^n_{t,u} = r_t P^n_{t,u}, \quad P^n_{t,0} = 1.
$$

We guess the solution form of the PDE as follows:

$$
P^n_{t,u} = \exp(X_tA^n u X_t + (b^n u)' X_t + c^n u).
$$

Substituting this into the PDE, we obtain the system of ordinary differential equations (ODEs).

$$
\begin{align*}
-\hat{A}^n_u - 2K^Q X_t A^n u + 2A^n u \Sigma_X \Sigma' X_t A^n u - \Psi &= 0, \\
-(b^n u)' + 2\theta^Q (K^Q)' A^n u - (b^n u)' K^Q + 2(b^n u)' \Sigma X' A^n u &= 0, \\
-c^n u + (K^Q (\theta^Q)' b^n u + \text{Tr} (\Sigma_X \Sigma' (A^n u + \frac{1}{2} b^n u (b^n u)')) &= 0,
\end{align*}
$$

where the boundary conditions are $A^n_0 = 0$, $b^n_0 = 0$ and $c^n_0 = 0$ and $\hat{A}^n, b^n$ and $c^n$ represent derivatives of $A^n, b^n$ and $c^n$ with respect to the variable $u$.

#### 3.2. Bond pricing in the abnormal period, the QE period

In this subsection, we derive a bond pricing formula in the case where $t \leq \tau$. This corresponds to the period when the Quantitative easing policy is conducted.

Firstly, we derive the bond pricing formula in the case where $t \leq T \leq \tau$. The zero coupon bond price during the period is denoted by $P^a_{t,u}$.

Infinitesimal generator of $X_t$ and $y^\tau_t$ for $t < \tau$ is provided as

$$
\mathcal{D}_t^a = (K^Q X_t (\theta^Q - X_t))' \frac{\partial}{\partial X_t} + \frac{1}{2} \text{Tr} \left( \Sigma_X \Sigma' X_t \frac{\partial^2}{\partial X_t \partial X_t'} \right) - \frac{y^\tau_t}{\tau - t} \frac{\partial}{\partial y^\tau_t} + \frac{1}{2} \sigma_y^2 \frac{\partial^2}{\partial y^\tau_t^2}.
$$

Applying the Feynman-Kac theorem to $P^a_{t,u}$, we obtain the following PDE:

$$
\left[ \frac{\partial}{\partial t} + \mathcal{D}_t^a \right] P^a_{t,u} = r_t P^a_{t,u}, \quad P^a_{t,0} = 1.
$$
A guess the solution of this PDE is provided as
\[ P_{t,u}^{a,1} = \exp(X'_u A^{a,1}_u X_t + (b^{a,1}_u)' X_t + c^{a,1}_u + d^{a,1}_u y_t). \]
Substituting the above solution form into the PDE, we obtain the ODEs.
\[
\begin{align*}
-\dot{A}^{a,1}_u - 2K'_u A^{a,1}_u + 2A^{a,1}_u \Sigma^X_u A^{a,1}_u - \Psi &= 0 \\
-(b^{a,1}_u)' + 2 \theta^Q (K^Q)' A^{a,1}_u - (b^{a,1}_u)' K^Q + 2(b^{a,1}_u)' \Sigma^X_u A^{a,1}_u &= 0 \\
-c^{a,1}_u + (K^Q \theta^Q)' A^{a,1}_u + \text{Tr}(\Sigma^X_u \Sigma^X_u (A^{a,1}_u + \frac{1}{2} b^{a,1}_u (b^{a,1}_u)')) + \frac{1}{2} \sigma^2 (d^{a,1}_u)^2 &= 0 \\
-d^{a,1}_u = 1 &= - \frac{w}{2(u-w)}.
\end{align*}
\]
where boundary conditions are \( A^{a,1}_0 = 0, b^{a,1}_0 = 0, c^{a,1}_0 = 0, \) and \( d^{a,1}_0 = 0. \) Setting \( w = T - \tau, \) the solution of the equation for \( d^{a,1}_0 \) (that is called d’Alembert’s equation) is given by \( d^{a,1}_0 = - \frac{w(u-2w)}{2(u-w)}. \) Since \( c^{a,1}_u \)
and \( d^{a,1}_u \) depend on \( u, \) we denote them by \( c^{a,1}_u, w, \)
and \( d^{a,1}_u, w, \).
Next, we derive the bond pricing formula in the case of \( t \leq \tau \leq T. \) We set \( v = \tau - t \) and \( w = T - \tau. \)
We denote the zero coupon bond price \( P_{t,v\tau}^{a,2} \) during the period by \( P_{t,v\tau}^{a,2} \).
By applying the Feynman-Kac theorem to \( P_{t,v\tau}^{a,2} \),
\[
\frac{\partial}{\partial t} P_{t,v\tau}^{a,2} - \mathcal{G} P_{t,v\tau}^{a,2} = r_t P_{t,v\tau}^{a,2}, \quad P_{t,0w}^{a,2} = P_{t,w}^{a,2},
\]
A guess of the solution of this equation:
\[ P_{t,v\tau}^{a,2} = \exp(X'_u A^{a,2}_u X_t + (b^{a,2}_u)' X_t + c^{a,2}_u + d^{a,2}_u y_t). \]
By substituting this solution form into the above PDE,
\[
\begin{align*}
-\dot{A}^{a,2}_u - 2K'_u A^{a,2}_u + 2A^{a,2}_u \Sigma^X_u A^{a,2}_u - \Psi &= 0 \\
-(b^{a,2}_u)' + 2 \theta^Q (K^Q)' A^{a,2}_u - (b^{a,2}_u)' K^Q + 2(b^{a,2}_u)' \Sigma^X_u A^{a,2}_u &= 0 \\
-c^{a,2}_u + (K^Q \theta^Q)' A^{a,2}_u + \text{Tr}(\Sigma^X_u \Sigma^X_u (A^{a,2}_u + \frac{1}{2} b^{a,2}_u (b^{a,2}_u)')) + \frac{1}{2} \sigma^2 (d^{a,2}_u)^2 &= 0 \\
-d^{a,2}_u + c^{a,2}_u + 1 &= 0
\end{align*}
\]
where boundary conditions \( v = 0 \) are \( A^{a,2}_0 = A^{a,2}_0, b^{a,2}_0 = b^{a,2}_u, c^{a,2}_0 = c^{a,2}_u, \)
and \( d^{a,2}_u \) with respect to the variable \( u \) for a constant \( w. \)
We haven’t had the boundary condition of \( d^{a,2}_u \) yet.
\[
P_{t,v\tau}^{a,2} = E \left[ \exp \left( - \int_{\tau}^{T} r_s ds \right) P_{T,v\tau}^{a,2} \big| F_{\tau} \right] = E \left[ \exp \left( - \int_{\tau}^{T} (X'_s \Psi X_s + y_t')ds \right) P_{T,v\tau}^{a,2} \big| F_{\tau} \right]
\]
The first expectation of r.h.s of the above equation depends on \( y_t \), and the second expectation does not depend on that. Hence, the following equation holds.
\[
E \left[ \exp \left( - \int_{\tau}^{T} y_t' ds \right) \big| F_{\tau} \right] = \exp(d^{a,2}_v y_t).
\]
The l.h.s of the above equation does not depend on \( w; \) therefore, \( d^{a,2}_v = d^{a,2}_v(0) \). Recalling that when \( \tau = T \) in the discussion in the case of \( t \leq \tau \leq T \)
\[
d^{a,2}_v = d^{a,2}_0 = d^{a,1}_0 = - \frac{1}{2} v.
\]
4. Bond pricing in the case where \( \tau \) is random. In this section, we derive the bond pricing formula in the case where \( \tau \) is random. The Brownian bridge with a random terminal time was studied by Bedini et al [1]. We use some results obtained in Bedini et al [1].
Let \( (C, C) \) be the space of continuous real-valued functions on \( \mathbb{R} \) endowed with the \( \sigma \)-algebra generated by the canonical process. We define \( y_t(\omega) \) as \( y_t(\omega) = y_t^x(\omega) \). Now we can prove that the map \( y : (\Omega, \mathcal{F}) \to (C, C) \) is measurable. The following corollary and lemma play roles in deriving zero coupon bond pricing formula.
Corollary 4.1. τ is assumed to be independent of \( W_{t^1}^Q \) and \( W_{t^2}^Q \). If \( h : ((0, +\infty) \times C, \mathcal{B}((0, +\infty)) \otimes \mathbb{R}, \mathcal{B}(\mathbb{R})) \) is a measurable function such that \( E[|h(\tau, y)|] < +\infty \), then \( E[h(\tau, y)|\sigma(\tau)](\omega) = E[h(\tau, y)]|_{\tau=\tau(\omega)}, Q - a.s. \)

Lemma 4.2. The density of \( y_t^* \) is provided as follows:

\[
\phi_t(r, y) = \frac{r}{2\pi t(r - t)\sigma_y^2} \exp\left(-\frac{y^2r}{2t(r - t)\sigma_y^2}\right).
\]

Let \( F(x) \) be the distribution function of \( \tau, Q(\tau \leq x) \) and \( f(x) \) be the density function of \( \tau \) under \( Q \). We define \( G(t) \) as

\[
G(t) = \int_t^\infty \phi_t(s, y_t^*)dF(s).
\]

We define the filtration \( \mathcal{H} := \sigma(1_{\tau \leq s} : 0 \leq s \leq t) \) and \( \mathcal{G}_t := \mathcal{F}_t \vee \mathcal{H}_t \). When \( \tau \) is independent from \( W_{t^1}^Q \) and \( W_{t^2}^Q \), the zero coupon bond prices are calculated as follows: This corollary implies that if \( \tau \) is independent from \( W_{t^1}^Q \) and \( W_{t^2}^Q \), then zero coupon bond price is obtained by taking expectations of bond prices with a constant \( \tau \) over the \( \tau \)'s probability distribution.

\[
P_{t, T-t} = E\left[ \exp\left(-\int_t^T r_s ds\right) | \mathcal{G}_t \right] = E\left[ 1_{\{\tau < t\}} \exp\left(-\int_t^T r_s ds\right) | \mathcal{G}_t \right] + E\left[ 1_{\{t \leq \tau \leq T\}} \exp\left(-\int_t^T r_s ds\right) | \mathcal{G}_t \right] + E\left[ 1_{\{T < \tau\}} \exp\left(-\int_t^T r_s ds\right) | \mathcal{G}_t \right]. \tag{1}
\]

The first term of the r.h.s of equation (1) is calculated as follows:

\[
E\left[ 1_{\{\tau < t\}} \exp\left(-\int_t^T r_s ds\right) | \mathcal{G}_t \right] = 1_{\{\tau < t\}} P_{t, T-t}^{\text{up}}
\]

The third term of the r.h.s of equation (1) is calculated as follows:

\[
E\left[ 1_{\{T < \tau\}} \exp\left(-\int_t^T r_s ds\right) | \mathcal{G}_t \right] = 1_{\{T < \tau\}} \frac{G(T)}{G(t)} P_{t, T-t}^{\text{up}}
\]

The second term of the r.h.s of equation (1) is calculated as follows:

\[
E\left[ 1_{\{t \leq \tau \leq T\}} \exp\left(-\int_t^T r_s ds\right) | \mathcal{G}_t \right] = \frac{1_{\{t < \tau\}}}{G(t)} \int_t^T P_{t, s-t}^{\text{up}} \phi_t(s, y_t^*) f(s) ds.
\]

5. Empirical Analysis. We conduct an empirical analysis using Japanese yield curve databased on our proposed model. I will indicate some results in my presentation.

References


Arbitrage-free option prices interpolation by stochastic collocation. Summary and applications

FABIEN LE FLOC’H
DELFt INSTITUTE OF APPLIED MATHEMATICS, TU DELFT, DELFT, THE NETHERLANDS
EMAIL: F.L.Y.LeFloch@tudelft.nl

CORNELIS W. OOSTERLEE
CWI-CENTRUM WISKUNDE & INFORMATICA, AMSTERDAM, THE NETHERLANDS

Abstract. This paper summarizes the arbitrage-free stochastic collocation technique. We review its application for market option price interpolation, through a specific isotonic polynomial or a monotonic spline representation, as well as in the context of the Heston collocated local volatility model. We present how to perform the calibration of this model efficiently, explore the preservation of the martingale property and practical Monte-Carlo and partial differential equation discretization schemes.

1. Summary. The financial markets provide option prices for a discrete set of strike prices and maturity dates. In order to price over-the-counter vanilla options with different strikes, or to hedge complex derivatives with vanilla options, it is useful to have a continuous arbitrage-free representation of the option prices, or equivalently of their implied volatilities. For example, the variance swap replication of Carr and Madan consists in integrating a specific function over a continuum of vanilla put and call option prices [7, 6]. More generally, [4] have shown that any path-independent claim can be valued by integrating over the probability density implied by market option prices. An arbitrage-free representation is also particularly important for the Dupire local volatility model [9], where arbitrage will translate to a negative local variance. In this paper, we describe a new technique to interpolate the market option prices in an arbitrage-free manner.

A rudimentary, but popular representation is to interpolate market implied volatilities with a cubic spline across option strikes. Unfortunately this may not be arbitrage-free as it does not preserve the convexity of option prices in general. A typical convex interpolation of the call option prices by quadratic splines or rational splines is also not satisfactory in general since it may generate unrealistic oscillations in the corresponding implied volatilities, as evidenced in [18]. [19] designs an arbitrage-free interpolation of the call prices, which however requires convex input quotes, employs two embedded non-linear minimizations, and it is not proven that the algorithm for the interpolation function of class $C^2$ converges. In reality it is often not desirable to strictly interpolate option prices as those fluctuate within a bid-ask spread. Interpolation will lead to a noisy estimate of the probability density (which corresponds to the second derivative of the undiscounted call option price).

More recently, [3] have proposed to calibrate the discrete piecewise constant local volatility corresponding to a single-step finite difference discretization of the forward Dupire equation. In their representation of the local volatility, the authors use as many constants as the number of market option strikes for an optimal fit. It is thus sometimes considered to be “non-parametric”. Their technique works well but often yields a noisy probability density estimate, as the prices are typically over-fitted. Furthermore the output of their technique is a discrete set of option prices, which, while relatively dense, must still be interpolated carefully to obtain the price of options whose strike falls in between nodes.

An alternative is to rely on a richer underlying stochastic model, which allows for some flexibility in the implied volatility smile, such as the Heston or SABR stochastic volatility models. While semi-analytic representations of the call option price exist for the Heston model [16], the model itself does not allow to represent short maturity smiles accurately. The SABR model is better suited for this, but has only closed-form approximations for the call option price, which can lead to arbitrage [14, 13]. Furthermore
the small number of parameters of such models is not always adequate to represent volatility surfaces, particularly for liquid equity or equity index options.

[12] use stochastic collocation to fix the Hagan SABR approximation formula defects and produce arbitrage-free option prices starting from the Hagan SABR formula. Collocation methods are commonly used to solve ordinary or partial differential equations [25]. The underlying principle is to solve the equations in a specific finite dimensional space of solutions, such as polynomials up to a certain degree. In contrast, the stochastic collocation method [26] consists in mapping a physical random variable $Y$ to a point $X$ in an artificial stochastic space. Collocation points $x_i$ are used to approximate the function mapping $X$ to $Y$, $F_X^{-1} \circ F_Y$, typically by a polynomial, where $F_X$, $F_Y$ are respectively the cumulative distribution functions (CDF) of $X$ and $Y$. Thus only a small number of inversions of $Y$ (and evaluations of $F_Y$) are used. This allows the problem to be solved in the “cheaper” artificial space. Using a Gaussian distribution for $X$, European option prices are obtained by a simple closed-form formula, for any strike. A necessary condition for the prices to be free of arbitrage, is for the mapping to be monotonic, otherwise the collocated CDF may decrease. This is, a priori, not guaranteed with the approach of [12], where a Lagrange polynomial on Gauss-Hermite nodes is used.

2. Applications. [22] explore how to calibrate the stochastic collocation polynomial directly to market prices, without going through an intermediate model. A specific isotonic parameterization is used to ensure the monotonicity of the collocation polynomial as well as the conservation of the zero-th and first moments transparently during the optimization, guaranteeing the absence of arbitrage.

The isotonic polynomial stochastic collocation leads to a smooth, implied probability density, without any artificial peak, even with high degrees of the collocation polynomial. The technique may also be applied to interest rate derivatives, as it leads to a closed-form formula for CMS convexity adjustments, which can thus be easily calibrated jointly with interest rate swaptions.

The stochastic collocation technique is of particular interest to the richer collocated local volatility (CLV) model, which allows to price exotic options through Monte-Carlo or finite difference methods [11]. In this model, the financial asset $S$ follows

$$S(t) = g(t, X(t)),$$

$$dX(t) = (r(t) - q(t))X(t)dt + \sqrt{V(t)}X(t) \, dW_X(t),$$

$$dV(t) = \kappa(\theta - V(t)) + \sigma \sqrt{V(t)} \, dW_V(t),$$

with $W_X$ and $W_V$ being two Brownian motions with correlation $\rho$, $r, q$ the instantaneous growth and dividend rates, and $g(t, x)$ a collocation function. According to equations (2) and (3), the driver process $X$ follows the Heston stochastic volatility model [16]. The collocation function $g$ is typically a polynomial or a spline, which will be calibrated to the market using the stochastic collocation technique. Robust and efficient pricing of vanilla options is obtained by applying the adaptive Filon quadrature of [21] on a specific payoff.

The use of a polynomial may however be problematic as the (lognormal) Heston model suffers from moment explosions [2]. The moment $\mathbb{E}[S(t)^m]$ for $m \geq 1$ may become infinite after some finite time. In those cases, which are quite common when the Heston model is calibrated to the market, the first moment of the Heston-CLV model is not defined. While it is always possible to constrain the Heston parameters to a region where the moments are well defined, such a constraint places a severe limitation on the Heston model quality of fit (Figures 1(a) and 1(b)). This has potential consequences on the martingale conservation properties of the Heston-CLV model.

A known deficiency of Markov functional models is that they do not respect the martingale property [5, 17]. The Heston-CLV model is similar to a Markov functional model: the collocation function only captures the terminal distribution at each maturity date and is then applied at different dates independently, without taking into account any joint distribution between the dates. Yet, the driver is a calibrated Heston model, whose joint distribution is close to the market distribution. We may thus expect the collocation function to provide a second order correction and the martingale property to be reasonably well preserved. This is confirmed on our example calibration towards SX5E options as of
(a) Well calibrated Heston model, but with exploding second moment.

(b) Heston model with finite fifth moment.

Figure 1: Black-Scholes volatility implied by the Heston model, and the Heston-CLV model for the options of maturity 1 month, calibrated to TSLA options of maturities 1, 7 and 19 months as of June 18, 2018. The strike indicated is relative towards the forward price.

(a) Heston model.

(b) Heston-CLV model.

February 26, 2016 (Table 1 and Figure 2(b)). The measured drift is an order of magnitude smaller than with one-dimensional Markov functional models.
Table 1: Error in the price of a future contract on performance, expressed in %, for each pairs of distinct maturities, obtained by a Monte-Carlo simulation using 1 million paths and 32 time-steps per year, when the Heston-CLV model is calibrated to SX5E options as of February 26, 2016. The Monte-carlo standard error is less than 0.09% and the Heston parameters are $V_0 = 0.133$, $\kappa = 0.350$, $\theta = 0.321$, $\sigma = 1.388$, $\rho = -0.630$.

<table>
<thead>
<tr>
<th>Maturity</th>
<th>0.04</th>
<th>0.08</th>
<th>0.16</th>
<th>0.25</th>
<th>0.33</th>
<th>0.41</th>
<th>0.50</th>
<th>0.75</th>
<th>1.00</th>
<th>1.50</th>
<th>2.00</th>
<th>3.00</th>
<th>4.00</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.08</td>
<td>-0.02</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.16</td>
<td>-0.01</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.25</td>
<td>-0.02</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.33</td>
<td>-0.06</td>
<td>-0.05</td>
<td>-0.02</td>
<td>0.02</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.41</td>
<td>-0.06</td>
<td>-0.02</td>
<td>-0.03</td>
<td>0.04</td>
<td>0.01</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.50</td>
<td>-0.05</td>
<td>-0.06</td>
<td>-0.02</td>
<td>0.06</td>
<td>0.08</td>
<td>0.09</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.55</td>
<td>-0.05</td>
<td>-0.06</td>
<td>-0.02</td>
<td>0.06</td>
<td>0.08</td>
<td>0.09</td>
<td>0.13</td>
<td>0.03</td>
<td>0.03</td>
<td>0.03</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.00</td>
<td>-0.13</td>
<td>-0.10</td>
<td>-0.13</td>
<td>-0.09</td>
<td>-0.04</td>
<td>-0.10</td>
<td>-0.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.50</td>
<td>-0.08</td>
<td>-0.10</td>
<td>-0.14</td>
<td>-0.18</td>
<td>-0.22</td>
<td>-0.21</td>
<td>-0.28</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.00</td>
<td>-0.05</td>
<td>-0.20</td>
<td>-0.03</td>
<td>-0.12</td>
<td>-0.08</td>
<td>-0.09</td>
<td>-0.20</td>
<td>-0.22</td>
<td>-0.31</td>
<td>-0.14</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.00</td>
<td>0.01</td>
<td>-0.02</td>
<td>0.03</td>
<td>0.07</td>
<td>0.03</td>
<td>0.04</td>
<td>-0.01</td>
<td>-0.04</td>
<td>-0.20</td>
<td>-0.06</td>
<td>0.16</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4.00</td>
<td>0.01</td>
<td>-0.04</td>
<td>0.10</td>
<td>0.14</td>
<td>0.12</td>
<td>0.14</td>
<td>0.06</td>
<td>0.04</td>
<td>-0.13</td>
<td>-0.10</td>
<td>0.11</td>
<td>-0.11</td>
<td>-0.30</td>
</tr>
<tr>
<td>5.01</td>
<td>0.04</td>
<td>0.13</td>
<td>0.09</td>
<td>0.10</td>
<td>0.16</td>
<td>0.13</td>
<td>0.07</td>
<td>0.10</td>
<td>-0.06</td>
<td>0.00</td>
<td>0.05</td>
<td>-0.11</td>
<td>0.31</td>
</tr>
</tbody>
</table>

Another drawback of the use of polynomials in the stochastic collocation technique, independently of the Heston-CLV model, relates to their difficulties in capturing multi-modal distributions. Although, theoretically, as proven in [22], we can always find a polynomial to capture multi-modal distributions, such a polynomial may be required to be of a prohibitively high degree to match accurately the distribution, which renders the technique not practical.

Instead of collocating on a polynomial, [23] explore various ways to use a monotonic spline, including B-spline parametrizations which preserve the first moment exactly. This allows for a richer representation, with as many parameters as there are market option strikes. A direct consequence is the ability to capture more complex implied probability distributions such as multi-modal distributions. Overfitting is avoided by adding an appropriate regularization, whose optimal value may be found with the L-curve method of [15]. This is reminiscent of the penalized B-spline technique for volatility modelling of [8], where a B-spline parameterization of the Radon-Nikodym derivative of the underlying’s risk-neutral probability density with respect to a roughly calibrated base model is used. Concretely, Corlay’s method translates to an explicit probability density representation where the probability density is a spline multiplied by a base probability density function, such as the lognormal or normal probability density function. Corlay’s technique however limits the implied volatility shapes allowed, and often requires the use of a more elaborate base probability density function, such as the one stemming from the SVI parameterization of [10], to properly fit the market in practice. The stochastic collocation on a spline is more flexible and can fit the market very well when collocating to a simple Gaussian variable. Furthermore, with a linear extrapolation, the Heston-CLV model is applicable without any constraint on the Heston parameters.

The Monte-Carlo simulation of the Heston-CLV model is straightforward: we use a good discretization scheme for the Heston model in order to discretize the $X$ process; then we obtain the value of $S$ on each path by applying directly the collocation function to each path of $X$ (equation (1)). Unlike the case of the stochastic local volatility model, there is no need to use very small time-steps. The discretization error is entirely due to the Heston process discretization. The quadratic exponential (QE) scheme of [1] is widely used to discretize the Heston process. Its convergence properties are however not known (on a concrete example, we measure a convergence order between 1.3 and 1.7). Instead, we present a minor modification of the DVSS2 scheme of [24], which has a proven second order convergence, and stays computationally efficient.

We also explore different ways of solving the Heston-CLV partial differential equation (PDE): directly through the $L$-stable second-order finite difference scheme of [20], as well as with various alternating direction implicit (ADI) methods. All schemes explored may lead to oscillations, particularly visible in the Gamma, except the Lawson-Swayne scheme. The oscillations are attenuated by implicit Euler damping steps, but not removed. It is found that the Peaceman-Rachford scheme with Strang splitting damping steps is the most efficient, when applied to an orthogonalized version of the PDE.
References


Introducing Two Mixing Fractions to a Lognormal LSV Model

GEOFFREY LEE
DATA61 PRIVATE BAG 10 CLAYTON SOUTH VIC 3169 AUSTRALIA
EMAIL: Geoffrey.Lee@data61.csiro.au

ZILI ZHU
DATA61 PRIVATE BAG 10 CLAYTON SOUTH VIC 3169 AUSTRALIA
EMAIL: Zili.Zhu@data61.csiro.au

BOWIE OWENS
DATA61 PRIVATE BAG 10 CLAYTON SOUTH VIC 3169 AUSTRALIA
EMAIL: Bowie.Owens@data61.csiro.au

Abstract. Current local-stochastic volatility (LSV) models use a single parameter, termed the Mixing Fraction, to calibrate the LSV model to traded exotic prices as well as vanilla options. This single parameter has been applied to both the volatility of volatility parameter and the correlation between spot and volatility of the original stochastic volatility model. In this paper, we introduce two independent mixing fractions to separately apply impact to the volatility of volatility and the correlation. We observe significant improvement in calibration in accuracy to One Touch option prices with the introduction of a second mixing fraction.

1. Introduction. The local stochastic volatility (LSV) model is viewed in the industry as the state of the art model for pricing exotic options, such as barrier options and accumulators. Various forms of the LSV model have been implemented [1, 5, 6, 7, 8, 10, 11]. This work utilises the log-normal LSV model presented in [8]. The log-normal process for modelling the underlying volatility is a more accurate representation of market dynamics, due to zero probability for the volatility process to reach zero, unlike for example the Heston model when the Feller condition is often violated.

The mimicking theorem [4] ensures that LSV models are calibrated to reproduce vanilla option prices. However, for LSV models to be useful in practice, they must also replicate traded exotic prices. To achieve this objective, an extra parameter is typically introduced to enable the LSV model to also be calibrated to available traded exotic prices. This parameter typically multiplies the volatility of volatility parameter, and can be seen as controlling the stochasticity of the stochastic model, thus is viewed as representing a mixing of the local volatility and stochastic volatility. The parameter is thus termed a mixing fraction. Typically taking values between 0 and 1, this mixing fraction has also been applied to the correlation parameter [8, 9, 10].

However, applying a single mixing fraction to both stochastic parameters may have deleterious effects on calibration; decreasing the value of the volatility of volatility may improve calibration accuracy on its own, but when the correlation is also decreased, this improvement may be counteracted. In this work, we will present a modification to the log-normal LSV model by introducing a second mixing fraction to control the effect of correlation independently to that of the volatility of volatility parameter. We calibrate this modified log-normal local-stochastic volatility model to both vanilla and exotic prices available on the market, and we present numerical results to illustrate the improvement of pricing accuracy.

2. Model. The log-normal local-stochastic volatility (LSV) model is assumed to follow log-normal like dynamics for both the spot price \( S_t \) and the stochastic volatility \( \sigma_t \) as:

\[
\begin{align*}
  dS_t &= [r(t)]S_t dt + L(S_t, t)\sigma_t S_t dW^1_t, \quad S_0 = s, \\
  d\sigma_t &= \kappa(\theta - \sigma_t)dt + \eta\lambda\sigma_t dW^2_t, \quad \sigma_0 = v, \\
  \mathbb{E}[dW^1_t, dW^2_t] &= \eta dt.
\end{align*}
\]  

(1)
where \( r_d(t) \) is domestic interest rate and \( r_f(t) \) is foreign interest rate in the context of FX markets, both of which are assumed to be of term structure. We will denote \( r(t) = r_d(t) - r_f(t) \) in the remainder of this paper. We also assume that the stochastic parameters (\( \kappa, \theta, \lambda, \) and \( \rho \)) in the log-normal LSV model have term structures. Here \( L(S_t, t) \) is called the leverage function, which is numerically calibrated to the market data. \( L(S_t, t) \) can be viewed as representing the weight of local volatility.

To calibrate to market prices of exotic options, a mixing fraction \( \eta \) has conventionally been introduced, which acts on the vol of vol and correlation, such as in [8,9,10], and expressed by Equation 1. In this paper, we now introduce two mixing fractions, \( \eta_L \) and \( \eta_p \), and we obtain the following dynamics:

\[
\begin{align*}
\text{d}S_t &= [r(t)]S_t \text{d}t + L(S_t, t)\sigma_t S_t \text{d}W^1_t, \quad S_0 = s, \\
\text{d}\sigma_t &= \kappa(\theta - \sigma_t) \text{d}t + \eta_L \lambda \sigma_t \text{d}W^2_t, \quad \sigma_0 = v, \\
E[\text{d}W^1_t \cdot \text{d}W^2_t] &= \eta_p \rho \text{d}t.
\end{align*}
\]

\[\text{(2)}\]

2.1. Calibration of the log-normal LSV model. To calibrate the model, we first need to generate the local volatility surface \( \sigma_{LV}(S, t) \) from market input data on implied volatilities across different strikes and tenors, by using the local volatility formula [2,3]. The parameters of the plain stochastic volatility model are then calibrated to the same market data using a closed form expression for European puts under the log-normal stochastic volatility model. For more details on this calibration process, see [8].

At this stage, both local and stochastic components of the model are calibrated to the market implied volatility surface. The final stage of LSV calibration is to calibrate the term structure of the two mixing fractions to ensure the model matches chosen traded exotic options and compute the leverage function by solving a Fokker-Planck equation to find the transition probability density, \( p(S, V, t) \). We refer readers to [8,11] for a more detailed explanation, but will briefly outline the procedure below.

Using the mimicking theorem presented in [4], we can write the leverage function as

\[
L(x, t) = \frac{\sigma_{LV}(x, t)}{E[V|S_t = x]} = \sigma_{LV}(x, t) \left( \frac{\int_{R^+} p(x, V, t) dV}{\int_{R^+} V p(x, V, t) dV} \right).
\]

We can therefore obtain the leverage function at each time \( t \) by computing the integrals in Equation 3.

The calibration of the mixing fraction values for each tenor requires pricing exotic options, and minimising the root mean square error to their market prices. We use the Golden Section Search algorithm for calibrating a single mixing fraction, and the COntstrained Optimisation BY Linear Approximation (COBYLA) algorithm for calibrating two mixing fractions.

We solve both the Fokker-Planck equation (a PDE) and the backward partial-differentiation-equation (PDE) for pricing options, with the given initial and boundary conditions and also using an Alternative-Directional Implicit (ADI) scheme for discretisation; for more detailed information please see [8,11]. Generally, we impose zero second derivative condition and one-sided finite difference for first derivative. When there is a boundary condition for different option types, we normally use central finite difference; or one-sided difference accordingly (e.g., barrier options). It should be noted that the stochastic parameters are multiplied by the mixing fraction values in both PDEs.

3. Numerical Results. We use a traded market dataset to test the impact of using the two-parameter mixing fraction. The dataset contains 7 tenors of implied volatilities for ATM, 10% and 25% deltas for both vanilla calls and puts. We also have the domestic and foreign yields available at each tenor. We calibrate to two One Touch options at each tenor, one with a trigger below the spot and one with a trigger above the spot. The bounds for the first single mixing fraction \( \eta_L \) are [0,1,2], while the bounds for \( \eta_p \) are [\(-1/\rho, 1/\rho\)].

The objective error function when calibrating mixing fractions to exotic options prices at a single tenor is presented in Figure 1. The optimisation scheme aims to minimise the root mean square error of the priced exotic prices, and in Figure 1, the darker regions represent better parameter choices. The black dots shown in the figure indicate the errors from the calibration procedure using a single mixing fraction. Figure 1 demonstrates that there is a clear improvement in the calibration results when two mixing fractions are utilised.
The calibrated values of the mixing fractions are presented in Figure 2 for all available market maturity tenors. When moving to two mixing fractions, the volatility of volatility component remains similar to the original single parameter mixing fraction case over time. However, the correlation component introduces a major change to the calibrated results, where the optimal value of $\eta_\rho$ becomes significantly negative for short-dated tenors. This results in the effective correlation $\eta_\rho \rho$ changing sign from the originally calibrated value.

These changes in the correlation component result in improved model calibration to the One Touch exotic options. We present the upper trigger One Touch prices from the calibrated model in Table 1 and the lower trigger results in Table 2. There is a significant reduction in price difference when two mixing fractions are used, this is particularly evident for short-dated maturities in the lower trigger results.
### Two Mixing Fractions

Table 1: Calibration results for One Touch exotic options with an upper trigger

<table>
<thead>
<tr>
<th>Maturity</th>
<th>Market Price (% of spot)</th>
<th>Price with 1 mixing fraction</th>
<th>Difference (%)</th>
<th>Price with 2 mixing fractions</th>
<th>Difference (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1w</td>
<td>25.15</td>
<td>24.91</td>
<td>-0.24</td>
<td>24.63</td>
<td>-0.09</td>
</tr>
<tr>
<td>2w</td>
<td>24.83</td>
<td>24.64</td>
<td>-0.19</td>
<td>24.71</td>
<td>-0.12</td>
</tr>
<tr>
<td>1m</td>
<td>34.91</td>
<td>34.98</td>
<td>0.07</td>
<td>34.95</td>
<td>0.04</td>
</tr>
<tr>
<td>3m</td>
<td>36.02</td>
<td>35.22</td>
<td>-0.20</td>
<td>35.21</td>
<td>0.19</td>
</tr>
<tr>
<td>6m</td>
<td>34.15</td>
<td>34.20</td>
<td>0.27</td>
<td>34.13</td>
<td>-0.01</td>
</tr>
<tr>
<td>1y</td>
<td>35.60</td>
<td>35.50</td>
<td>-0.00</td>
<td>35.50</td>
<td>0.00</td>
</tr>
</tbody>
</table>

Table 2: Calibration results for One Touch exotic options with a lower trigger

<table>
<thead>
<tr>
<th>Maturity</th>
<th>Market Price (% of spot)</th>
<th>Price with 1 mixing fraction</th>
<th>Difference (%)</th>
<th>Price with 2 mixing fractions</th>
<th>Difference (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1w</td>
<td>26.14</td>
<td>25.77</td>
<td>-0.07</td>
<td>26.15</td>
<td>-0.00</td>
</tr>
<tr>
<td>2w</td>
<td>26.56</td>
<td>27.28</td>
<td>0.72</td>
<td>26.59</td>
<td>0.03</td>
</tr>
<tr>
<td>1m</td>
<td>35.62</td>
<td>35.92</td>
<td>-0.27</td>
<td>35.92</td>
<td>0.30</td>
</tr>
<tr>
<td>3m</td>
<td>36.07</td>
<td>35.92</td>
<td>-0.15</td>
<td>35.92</td>
<td>0.00</td>
</tr>
<tr>
<td>6m</td>
<td>36.08</td>
<td>36.08</td>
<td>0.00</td>
<td>36.08</td>
<td>0.00</td>
</tr>
<tr>
<td>1y</td>
<td>35.83</td>
<td>35.83</td>
<td>0.00</td>
<td>35.83</td>
<td>0.00</td>
</tr>
<tr>
<td>2y</td>
<td>34.28</td>
<td>34.55</td>
<td>-0.27</td>
<td>34.50</td>
<td>-0.03</td>
</tr>
</tbody>
</table>

4. Conclusion. We have introduced two mixing fraction parameters into the calibration of the Log-normal LSV model to separate the effects of the volatility of volatility and correlation between spot and volatility. We observe improved calibration results over the use of a single mixing fraction, exotic prices from the two-mixing-fraction model are demonstrably closer to market prices, particularly for lower trigger One Touch options. Improved calibration to traded exotic prices may lead to better hedging performance and improved risk management.

References


Abstract. A real option on finite time horizon is considered under a regime-switching jump-diffusion model. The investor in the real option wants to decide an optimal investment time to maximize the discounted expectation of a payoff function. In this talk, the project value is evaluated by solving a partial integro-differential equation (PIDE) and it can be expressed as the closed-form solution. Then the objective function and the optimal investment time can be computed by using an operator splitting method. Numerical experiments are carried out to describe the various phenomena with the regime-switching process.
Continuous Time Markov Chain approximation of the Heston model

ÁLVARO LEITAO∗  
BGSMath and Department of Econometrics, University of Barcelona  
Email: leitao@ub.edu  
URL: https://alvaroleitao.github.io

LUIS ORTIZ-GRACIA†  
Department of Econometrics, University of Barcelona  
Email: luis.ortiz-gracia@ub.edu  
URL: https://sites.google.com/site/luorgr

JUSTIN L. KIRKBY  
IDology

Abstract. In this work an approximation of the Heston model is presented, which facilitates its numerical treatment, not only for simple problems (European options) but also for more sophisticated products (exotic options), in a unified framework. The presented approach is based on the so-called Continuous Time Markov Chain (CTMC) approximation applied to one of the equations in the Heston model, particularly the volatility dynamics.

1. From Heston model to CTMC-Heston model. The Heston model [1] is a stochastic volatility model defined by the following system of stochastic differential equations (SDEs),

\[
\begin{align*}
\text{d}S_t &= (r - d)S_t \text{d}t + \sqrt{v_t}S_t \text{d}W^1_t, \\
\text{d}v_t &= \eta(\theta - v_t) \text{d}t + \sigma_v \sqrt{v_t} \text{d}W^2_t,
\end{align*}
\]

with \(\text{d}W^1_t\) and \(\text{d}W^2_t\) to correlated Brownian motions, i.e. \(\text{d}W^1_t \text{d}W^2_t = \rho \text{d}t\). The stochastic volatility (or variance), \(v_t\), is driven by a CIR process, having a mean reverting component. The model parameters are therefore \(v_0, \eta, \theta, \sigma_v\) and \(\rho\).

1.1. CTMC approximation of the variance process. This approach is based on some developments presented in [6] and the references therein. Let's consider the log-return of the underlying \(X_t = \log \left( \frac{S_t}{S_0} \right)\) and an auxiliary process defined as \(\hat{X}_t = X_t - g(v_t, v_0)\), with \(g(v_t, v_0) = \frac{\rho}{\sigma_v} (v_t - v_0)\). Then, the Heston SDEs are re-written in such a way that the two equations become independent, i.e. there is no correlation between them. Thus, the log-transformed and non-correlated version of the Heston model reads

\[
\begin{align*}
\text{d}\hat{X}_t &= \left( \left( \frac{\rho \eta}{\sigma_v} - \frac{1}{2} \right)v_t + \left( r - d - \frac{\rho \theta}{\sigma_v} \right) \right) \text{d}t + \sqrt{(1 - \rho^2)}v_t \text{d}W^*_t, \\
\text{d}v_t &= \eta(\theta - v_t) \text{d}t + \sigma_v \sqrt{v_t} \text{d}W^2_t,
\end{align*}
\]

where now the Brownian motions \(\text{d}W^*_t\) and \(\text{d}W^2_t\) are uncorrelated.

Next, we approximate the variance process, \(v_t\), by \(\tilde{m}=\text{state CTMC, denoted by } v_{\alpha(t)}\), where the chain \(\alpha(t)\) transitions between states \(\{1, 2, \ldots, \tilde{m}\}\). Doing so, we can reformulate the decorrelated Heston model in Equation (2) and cast it in the formulation of the Regime-Switching (RS) processes as

\[
\text{d}\hat{X}_t = \zeta_{\alpha(t)} \text{d}t + \beta_{\alpha(t)} \text{d}W^*_t,
\]

∗Álvaro Leitao acknowledges financial support from the Spanish Ministry of Science and Innovation, through the María de Maeztu Programme for Units of Excellence in R&D (MDM-2014-0445).
†Luis Ortiz-Gracia acknowledges the Spanish Ministry of Economy and Competitiveness for funding under grants ECO2016-76203-C2-2 and MTM2016-76420-P (MINECO/FEDER, UE).
where
\[ \zeta_{\alpha(t)} = \left( \frac{\rho \eta}{\sigma_v} - \frac{1}{2} \right) v_{\alpha(t)} + \left( (r - d) - \rho \eta \theta \right) \beta_{\alpha(t)} = \sqrt{1 - \rho^2} v_{\alpha(t)}, \]
and \( \alpha(t) \) independent of \( W^*_t \).

The explicit representation of the CTMC-Heston model in the framework of RS processes turns to be very convenient. The availability of the characteristic function is particularly interesting. Assuming a variable \( Y \) driven by a RS dynamics, the characteristic function of \( Y_{\Delta t} \) for each state \( j = 1, \ldots, m_0 \) satisfies
\[ \mathbb{E} \left[ e^{i \xi Y_{\Delta t}} | \alpha(0) = j \right] = e^{ \mathcal{M}(\xi) \cdot \mathbf{e}_j }, \quad \mathcal{M}(\xi) := e^{\Delta t (\mathcal{Q} + \text{diag}(\psi_1(\xi), \ldots, \psi_m(\xi)))}, \]
being \( \mathbf{1} \in \mathbb{R}^{m_0 \times 1} \) a vector of ones and \( \mathbf{e}_j \in \mathbb{R}^{m_0 \times 1} \) a unit vector with a one in the \( j \)-th position. Matrix \( \mathcal{Q} = \{ q_{jk} \}_{m_0 \times m_0} \) is the so-called generator, satisfying that \( q_{jj} \leq 0 \) and \( q_{jk} \geq 0 \) if \( j \neq k \), and \( \sum_k q_{jk} = 0, \forall j \). The generator fully describes the transition dynamics \( \alpha(t) \). The function \( \psi_j(\xi) \) is the so-called characteristic exponent or Lévy symbol, which is available for RS models. In the context of the CTMC-Heston model, the characteristic exponent will be (with the notation above)
\[ \psi_j(\xi) = i \xi \zeta_j - \frac{1}{2} \xi^2 \beta_j^2. \]

What we aim to estimate is the characteristic function of \( X_{\Delta t} = \tilde{X}_{\Delta t} + g(v_{\Delta t}, v_0) \), i.e., the log-transformed return of the underlying process condition on each state of the volatility process \( j \). Thus,
\[ \mathbb{E} \left[ e^{i \xi X_{\Delta t}} | \alpha(0) = j, \alpha(\Delta t) = k \right] = \mathbb{E} \left[ e^{i \xi \tilde{X}_{\Delta t}} | \alpha(0) = j, \alpha(\Delta t) = k \right] \cdot e^{i \xi g(v_0, v_1)} = M_{kj}(\xi) \cdot e^{i \xi g(v_0, v_1)}. \quad (3) \]

In order to evaluate the characteristic function above, a generator \( \mathcal{Q} \) describing the transitions of \( \alpha(t) \) is required. Contrary to the RS model where the generator is an input of the model (it drives the changes in regimes), here we need the to construct the generator \( \mathcal{Q} \) according to the CTMC approximation of the Heston volatility process. Given a grid of points \( \nu = \{ v_1, v_2, \ldots, v_{m_0} \} \) with grid spacings \( h_i = v_{i+1} - v_i \) and assuming that \( v_{\alpha(t)} \) takes values on \( \nu \), we follow the formula proposed in [2] to define matrix \( \mathcal{Q} \). In Section 2.1, more insight about the optimal construction of \( \nu \) is provided.

2. Calibration of the CTMC-Heston model. Given the characteristic function in Equation 3, we use an efficient wavelet-based Fourier inversion technique called SWIFT method[4] to price financial options. SWIFT has several desirable advantages in a calibration procedure: error control given an scale of approximation, robustness in the parameters selection, performance efficiency and accuracy.

In this work, we focus on the calibration aspects related to the application of the CTMC approximation, that is, we assume the pure Heston model parameters to be given. Thus, there are two major components affecting the quality of the CTMC-Heston version of the volatility process: the length and the distribution of the points in the Markov chain, i.e. \( m_0 \) and the locations \( v_1, v_2, \ldots, v_{m_0} \), respectively. We will study the “optimal” selection of these two components. Let us start with the parameter \( m_0 \). It has been proved in [3] that \( v_{\alpha(t)} \) weakly converges to its continuous counterpart \( v_t \), so
\[ v_{\alpha(t)} \to v_t \text{ as } \max\{ |h_i| : i = 1, \ldots, m_0 - 1 \} \to 0, \quad (4) \]
recalling \( h_i = v_{i+1} - v_i \). An immediate conclusion from this result is that increasing the length of the chain does not necessarily ensures convergence, but it is also required to reduce the wider gap between grid points. So, as expected, bigger \( m_0 \) values cannot be analysed without studying the distribution of the newly added points, which must be smartly selected to ensure convergence in the approximations.

2.1. Grid selection. We will analyse several conceptually different approaches available in the literature, which can be classified into two main categories, based on state domain truncation and distribution inversion methodologies.

The truncation of the state space often relies on the availability of the moments and use them to select \( v_1 \) and \( v_{m_0} \).
As a most simple approach, an equally-spaced grid of points is determined by

\[ v_i = v_1 + (i - 1) \frac{v_{m_0} - v_1}{m_0 - 1}, \quad i = 2, \ldots, m_0 - 1. \]

More involved schemes can be introduced by considering non-uniform grids, aiming to better capture the relevant parts of the truncated state space and to improve the convergence rate. Thus, the algorithm proposed by Tavella and Randall in [5] reads

\[ v_i = v_0 + \bar{\alpha} \sinh \left( c_2 \frac{i}{m_0} + c_1 \left( 1 - \frac{i}{m_0} \right) \right) \]

with

\[ c_1 = \text{arc sinh} \left( \frac{v_1 - v_0}{\bar{\alpha}} \right), \quad \text{and} \quad c_2 = \text{arc sinh} \left( \frac{v_{m_0} - v_0}{\bar{\alpha}} \right), \]

where parameter \( \bar{\alpha} \) controls the non-uniformity of the grid. Mijatović and Pistorius [3] presented a slight modification of the Tavella and Randall algorithm based on the use of subgrids. The grid points are therefore chosen as follows,

\[ v_i = \begin{cases} 
  v_0 + \sinh \left( \frac{2 - i}{m_0} \text{arc sinh}(v_1 - v_0) \right), & i = 2, \ldots, \frac{m_0}{2}, \\
  v_0 + \sinh \left( \frac{i - \frac{m_0}{2}}{\text{arc sinh}(v_{m_0} - v_0)} \right), & i = \frac{m_0}{2} + 1, \ldots, m_0 - 1.
\end{cases} \]

Finally, the last approach takes advantage of the full available information to select the grid of discrete points. Since the distribution of the volatility process in the Heston model is known (\( v_t \) is driven by a CIR process), it seems natural to exploit this fact in order to optimally determine the location of the points (see [2]). Thus, we find that the grid points \( v_1, v_2, \ldots, v_{m_0} \) can be chosen by inversion as

\[ v_i = F_{v_t + \Delta t}^{-1} \left( \frac{2i - 1}{2m_0} \right), \]

where \( F_{\{\cdot\}} \) is the cumulative distribution function of the given random variable which, in our case, is

\[ v_t + \Delta t \sim \chi^2 \left( \frac{4\theta}{\sigma^2}, \frac{v_t + \Delta t e^{-\eta \Delta t}}{c_{\Delta t}} \right), \quad c_{\Delta t} = \frac{4\eta}{\sigma^2 (1 - e^{-\eta \Delta t})}, \]

with \( \chi^2(d, c) \) the non-central chi square distribution with \( d \) degrees of freedom and non-centrally parameter \( c \).

3. Experiments. Next, we perform an empirical study of the four proposed grid selection approaches in the context of the CTMC-Heston model. The experiments will consist of assessing the error in pricing European options, for which a reference price is easily obtained from the Heston characteristic function (available in closed-form).

First, in Figure 1a, given a model parameter setting, the grid points locations (\( m_0 = 30 \)) obtained by each of the approaches presented above are depicted. We can clearly identify the different distribution schemes among them.

Then, the convergence in terms of number of points, \( m_0 \), is studied. As mentioned, the convergence in the CTMC approximation of the variance process \( v_t \) is ensured when condition in Equation (4) is satisfied. All the grid selection methodologies presented here are in that situation, so a similar convergence behaviour should be translated to the option valuation results. Indeed, in Figure 1b we can observe a monotonic decreasing in the pricing error when \( m_0 \) is increased.

Finally, a real calibration test is carried out. We consider call options on Microsoft quoted in January 2019. As it is common in practice, we adjust the model parameters to fit the implied volatility observed in the market. Once the Heston parameters are determined, we aim to select \( m_0 \) in such a way that
the CTMC-Heston implied volatility approximates the ones by Heston up to some prescribed tolerance. Again, we analyse the four methodologies presented in Section 2.1 for variance process grid selection.

In Figure 2 the calibration outcome presented. Each bulleted curve corresponds to CTMC-Heston estimation with an increasing values of \( m_0 \). The tolerance is set to \( 10^{-3} \) and the fit between curves is measured by the mean relative error with respect to the Heston implied volatility along the strikes, \( K \).

We also employ a limit in the maximum chain length, set to \( m_0 = 200 \).

Several interesting preliminary lessons can be extracted from the previous experiments:

- The grid distribution proposed by Lo and Skindilias, despite more intuitive, provides poorer estimations when either \( v_0 \) or \( \theta \) are high. Possibly this because the state space is not sufficiently well covered. However, it becomes a good choice for lower values of both the initial and mean-reverting volatilities, as it is clearly observable in the calibration experiment with real market data.

- Although the uniform approach performs surprisingly well in the test with synthetic parameters, the real calibration experiment shows a pretty inaccurate estimations for options far from at-the-money strike, as the method is not able to even achieve the required tolerance, showing even a rotation over the at-the-money value.

- The two methodologies relying on moment-based domain truncation and non-uniform distribution perform similarly, being the one by Tavella and Randall the best choice in general.

- The method by Mijatovic and Pistorius, turns to be highly precise for the real data studied here, requiring very few Markov states to achieve the prescribed accuracy.

References


Figure 2: Market parameters: call option, $S_0 = 105.36$, $K = \{65, 70, \ldots, 150, 155\}$, $r = 0.0246$ and $T = 0.4986$. Heston parameters: $v_0 = 0.0906$, $\eta = 0.8549$, $\theta = 0.1379$, $\sigma_v = 0.9976$, $\rho = -0.6187$.


Calibration neural networks for financial models

Shuaiqiang Liu
TU Delft, The Netherlands
Email: s.liu-4@tudelft.nl

Cornelis W. Oosterlee
CWI, Amsterdam, The Netherlands
Email: c.w.oosterlee@cwi.nl

Abstract. A data-driven approach is proposed to calibrate advanced financial asset price models, namely CaNN (Calibration Neural Networks). Particularly, determining the optimal values of several model parameters is formulated as training hidden neurons within a machine learning framework, based on available financial derivative (option) prices. In the CaNN we distinguish a forward pass computation, in which we train off-line the weights of the ANN, valuing options under many different asset model parameter settings, from a backward pass, in which we evaluate the trained ANN-solver on-line, aiming to find the weights of all neurons in one unknown input layer, to achieve a robust and rapid overall calibration. The rapid on-line option pricing and implied volatility computations by ANNs, in combination with the use of a global optimization method provides a fast and reliable technique while avoiding, as much as possible, getting stuck in local minima. In this work, the parameters or implied information of several financial models are recovered quickly.
Some results on the Hybrid Heston-SLV Model With Jumps: Pricing, Hedging and Sensitivities. A Malliavin approach

ZORORO S. MAKUMBE
UAE UNIVERSITY, DEPARTMENT OF MATHEMATICAL SCIENCES, AL-AIN, P.O. BOX 17551, UNITED ARAB EMIRATES
EMAIL: zmakumbe@uaeu.ac.ae

YOUSSEF EL-KHATIB
UAE UNIVERSITY, DEPARTMENT OF MATHEMATICAL SCIENCES, AL-AIN, P.O. BOX 17551, UNITED ARAB EMIRATES
EMAIL: Youssef_Elkhatib@uaeu.ac.ae

JOSEP VIVES
UNIVERSITAT DE BARCELONA, DEPARTAMENT DE MATEMÀTIQUES I INFORMÀTICA, GRAN VIA 585, 08007 BARCELONA, SPAIN
EMAIL: josep.vives@ub.edu

Abstract. We consider a hybrid stochastic Local Volatility (SLV) model considered by [3], [7], and [5] among many researchers with the aim of incorporating the strengths of various models in particular the Heston model [9], Constant elasticity of Variance (CEV) model [4] and by extension the SABR model [8]. The pricing of a European options is investigated using Monte Carlo methods, two PDE approaches. We also apply functional Itô Calculus to hedge the portfolio and to compute sensitivities.

Keywords: European options, Numerical simulations, Monte Carlo method, Stochastic volatility Black and Scholes Formula.

1. Introduction. Though the Black-Scholes model [2] is popular due to its simplicity and also that it has a closed form solution it has many weaknesses. Stochastic volatility models are proposed in order to accommodate the stylised facts observed in practice.

2. Preamble.

Definition 2.1. (Lévy Process) Let \((\Omega, \mathcal{F}, \mathbb{P})\) where \(\mathbb{P} = (\mathcal{F}_t)_{0 \leq t \leq T}\) be a complete filtered probability space. A real valued \(\mathcal{F}_t\)-adapted process \(\eta(t)\) with \(\eta(0) = 0\) a.s. is called a Lévy process if \(\eta(t)\) is continuous in probability and has stationary and independent increments.

In order to introduce the Malliavin derivative we would like to discuss the Chaotic expansion of a stochastic process. To unify notation, we shall follow [10] as follows:

\[
U_i = \begin{cases} [0, T] & \text{when } i = 0 \\ [0, T] \times \mathbb{R} & \text{when } i = 1 \end{cases}
\]

\[
dQ_i = \begin{cases} dW & \text{when } i = 0 \\ \hat{N}(.,.) & \text{when } i = 1 \end{cases}
\]

\[
d(Q_i) = \begin{cases} d\lambda & \text{when } i = 0 \\ d\lambda \times d\nu & \text{when } i = 1. \end{cases}
\]

Also,

\[
G_{j_1, \ldots, j_n} = \left\{ (u_1^{j_1}, \ldots, u_n^{j_n}) \in \prod_{i=1}^n U_{j_i} : 0 < t_1 < \cdots < t_n < T \right\}.
\]
On the Jump Heston-SLV model

where $j_i = 0$ or $1$ for $i = 1, n$ and

$$u^l_k = \begin{cases} t_k & \text{when } l = 0 \\ (t_k, x) & \text{when } l = 1. \end{cases}$$

**Definition 2.2.**
1. Suppose $g_{j_1, \ldots, j_n} \in L^2(G_{j_1, \ldots, j_n})$ then

$$J_n^{(j_1, \ldots, j_n)}(g_{j_1, \ldots, j_n}) = \int_{G_{j_1, \ldots, j_n}} g_{j_1, j_n}(u_1^1, \ldots, u_n^j) dQ_{j_1}(u_1^1), \ldots, dQ_{j_n}(u_n^j)$$

is known as the n-fold iterated integral.

2. Further more if $\hat{L}^2(G_{j_1, \ldots, j_n}) \subseteq L^2(G_{j_1, \ldots, j_n})$ is the set of symmetric functions, then

$$I_n^{(j_1, \ldots, j_n)}(g_{j_1, \ldots, j_n}) = \int_{G_{j_1, \ldots, j_n}} g_{j_1, j_n}(u_1^1, \ldots, u_n^j) dQ_{j_1}(u_1^j), \ldots, dQ_{j_n}(u_n^j) = nJ_n^{(j_1, \ldots, j_n)}(g_{j_1, \ldots, j_n}).$$

**Theorem 2.3.** (Chaotic representation Property) Given a random variable $F \in L^2(\mathcal{F}_T, \mathbb{P})$, there exists a unique sequence of $(g_{j_1, \ldots, j_n})_{n=0}^\infty \subseteq L^2(G_{j_1, \ldots, j_n})$ such that

$$F = E[F] + \sum_{n=1}^\infty \sum_{j_1, \ldots, j_n=0,1} J_n^{(j_1, \ldots, j_n)}(g_{j_1, \ldots, j_n}).$$

Furthermore, we have the isometry

$$\|F\|_{L^2(F)}^2 = E[F]^2 + \sum_{n=1}^\infty \sum_{j_1, \ldots, j_n=0,1} \|J_n^{(j_1, \ldots, j_n)}(g_{j_1, \ldots, j_n})\|_{L^2(G_{j_1, \ldots, j_n})}^2.$$  

We define here directional derivatives with respect to the Wiener process ($D^{(1)}_t = D^W_t$) and Poisson random measure ($D^{(1)}_{t,z} = D^{G_{t,z}}_t$). Let $G_{j_1, \ldots, j_n}$ be the set $G_{j_1, \ldots, j_n}$ with the $k^{th}$ element deleted.

**Definition 2.4.** (Directional Derivative)
1. Let $g_{j_1, \ldots, j_n} \in L^2(G_{j_1, \ldots, j_n})$ and $l = 0, 1$. Then

$$D^l u_j J_n^{(j_1, \ldots, j_n)}(g_{j_1, \ldots, j_n}) = \sum_{j_1, \ldots, j_n=0,1} \mathbb{I}_{(j_1=0)}J_{n-1}^{(j_1, \ldots, j_n)}(g_{j_1, \ldots, j_n})(\cdots, u', \ldots) \mathbb{I}_{G_{j_1, \ldots, j_n}(t)}$$

is called the derivative of $J_n^{(j_1, \ldots, j_n)}(g_{j_1, \ldots, j_n})$ in the $l^{th}$ direction.

2. Let $\mathbb{D}^l$ be the space of all random variables in $L^2(\Omega)$ that are differentiable in the $l^{th}$ direction, then

$$\mathbb{D}^l = \left\{ F \in L^2(\Omega), F = E[F] + \sum_{n=1}^\infty \sum_{j_1, \ldots, j_n=0,1} J_n^{(j_1, \ldots, j_n)}(g_{j_1, \ldots, j_n}) : \right. \left. \sum_{n=1}^\infty \sum_{j_1, \ldots, j_n=0,1} \sum_{i=1}^n \mathbb{I}_{(j_i=l)} \int_{U_i} \|g_{j_1, \ldots, j_n}\|_{L^2(G_{j_1, \ldots, j_n})} dQ_t > (u^l) < \infty \right\}$$

3. Let $F \in \mathbb{D}^l$. Then the derivative in the $l^{th}$ direction is given as

$$D^l u_j F = \sum_{n=1}^\infty \sum_{j_1, \ldots, j_n=0,1} \sum_{i=1}^n \mathbb{I}_{(j_i=l)} J_{n-1}^{(j_1, \ldots, j_n)}(g_{j_1, \ldots, j_n})(\cdots, u', \ldots) \mathbb{I}_{G_{j_1, \ldots, j_n}(t)}.$$  

Another important theorem for our work is as follows:

**Theorem 2.5.** (General Clark-Ocone-Haussman Formula) Let $F \in \mathbb{D}^{(1)} \cap \mathbb{D}^{(1)}$. Then,

$$F = E[F] + \int_0^T E[D^{(1)}_{t,z} F_{t,z}^\circ] dW_t + \int_0^T E[D^{(1)}_{t,z} F_{t,z}] N(dt, dz).$$

**Definition 2.6.** Let $h(u)$ be a measurable stochastic process for all $u \in U$ such that for all $u$, $h(u)$ is $\mathcal{F}_T$-measurable and $\mathbb{E}[(\int_U h^2(u) dQ_u)] < \infty$ then $h(u) = \sum_{n=0}^\infty \sum_{j_1, \ldots, j_n=0,1} I_n^{(j_1, \ldots, j_n)}(g_{j_1, \ldots, j_n})$. We define the Skorohod integral (the adjoint operator of the Malliavin derivative) as

$$\delta(h) = \int_U h^2(u) dQ(t) = \sum_{n=0}^\infty \sum_{j_1, \ldots, j_n=0,1} I_n^{(j_1, \ldots, j_n)(j_{n+1})}(g_{j_1, \ldots, j_n}).$$

202
3. The Model. From now on we will assume that $$(\Omega, \mathcal{F}, P)$$ is the risk neutral probability space and let the asset price $$(S_t)_{t \geq 0}$$ satisfy the following:

$$dS_t = rS_t dt + \sqrt{\nu_t} S_t \left( dW_t + c \int_{\mathbb{R}_0} \gamma(t,z) \hat{N}(dt,dz) \right)$$  \hspace{1cm} (7)

$$S(0) = x > 0$$

$$dv_t = \kappa(\theta - v_t) dt + \sigma \sqrt{v_t} dB_t$$ \hspace{1cm} (8)

$$v(0) = y > 0.$$  

$$(W_t, B_t)_{t \in [0,T]}$$ is a 2 dimensional Brownian motion such that $$d\langle W, B \rangle = \rho dt$$ where $$\rho \in (-1,1)$$. And $$\hat{N}(t,A) = N(t,A) - \nu(A) dt$$ is the compensated random measure.

- Setting $$c = 0$$ makes the model into a continuous case Hybrid stochastic volatility model.
- Further, if $$\alpha = 1$$ and $$c = 0$$ we have the Heston stochastic volatility model.
- If $$c = 0$$, and the process $$Y_t = 1$$ this gives us the CEV models.
- If $$Y_t = 1$$ and $$\alpha = 1$$ we have the Bates model.
- If we let $$Y_t = 1$$, $$c = 0$$ and $$\alpha = 1$$ the model reduces to the classical Black-Scholes model.

3.1. Hedging.

Definition 3.1. A predictable process $$\Psi(t) = (\eta(t), \xi(t))$$ is called admissible if

1. $$\Psi(t)$$ is $$\mathcal{F}_t$$ adapted

2. $$\mathbb{E} \left[ \sum_{j=1}^{n} \int_0^T \Psi_j^2(t) (\sum_{j=1}^{n} \sigma_{i,j}^2 + \int_{\mathbb{R}_0} \gamma_{i,j}(t,z) \nu(dz) dt) \right] < \infty.$$

Also, let the set of all $$\mathcal{F}$$-admissible portfolios be denoted by $$A_{\mathcal{F}}$$

Proposition 3.2. Let $$V_t$$ be a self financing portfolio of a claim $$f(T, S_T) \in \mathcal{F}_T$$ given by

$$V_t = \eta_t A_t + \xi_t S_t.$$  \hspace{1cm} (9)

The minimal variance portfolio $$(\eta^*_t, \xi^*_t)$$ that minimises the hedging error

$$\min_{\xi_t \in A_{\mathcal{F}}} \mathbb{E} \left[ (\hat{V}_t - \Phi(T, S_T))^2 \right] = \mathbb{E} \left[ (\hat{V}_T^* - \Phi(T, S_T))^2 \right].$$

is given as

$$\xi^*_t = \frac{\sigma(S_t, Y_t) E[D_t^1 \Phi(S_T)|\mathcal{F}_t] + \int_{\mathbb{R}_0} \sigma(S_t, Y_t) E[D^1_t \Phi(S_T)|\mathcal{F}_t]}{e^{-rt} \sigma^2(S_t, Y_t) + \int_{\mathbb{R}_0} e^{-rt} \sigma^2(S_t, Y_t) \gamma^2(t,z) \nu(dz)}$$ \hspace{1cm} (10)

$$\eta^*_t = \frac{V - \xi_t S_t}{A_t}$$ \hspace{1cm} (11)

3.2. Price Sensitivities. The price of the option is given as

$$C = \mathbb{E}_t [e^{\int_0^T r_s ds} f(S(T)) | \mathcal{F}_t]$$

where $$r_t$$ is the interest rate, $$\mathbb{E}_t[.] = \mathbb{E}[.|\mathcal{F}_t]$$ and $$f(S(T)) = (S(T) - K)^+$$ is the pay-off. In our study we will take $$r_t = r$$ to be a constant for all $$t \in [0,T].$$

Greeks are a measure of the responsiveness of the risk neutral option price to changes in various parameters. We will consider a general parameter $$\varphi$$ then each first order Greek can be considered as a derivative as follows

$$\varphi = \frac{\partial}{\partial \varphi} E[\Phi(S_T)]$$

There are several approaches to the price sensitivities in literature. Classical techniques involve the finite difference approach and the Monte Carlo method which involve a high number of simulations and hence has a slow convergence rate. Other methods include pathwise method which requires a differentiable
payoff function and is not appropriate for complicated options like barrier or digital options. We shall
consider the Malliavin Calculus approach which has been found to reduce the number of computations
needed for the estimates to be made and hence has a much faster convergence rate. It has found wide
applications depending on the types of models to be analysed. [6] covers a process driven by jumps.

**Proposition 3.3.** The first order sensitivities of the plain vanilla option are given by

\[
\varrho = \frac{\partial}{\partial \varphi} \mathbb{E}[e^{-rT} \Phi(S_T)]
\]

(12)

\[
e^{-rT} \mathbb{E}[\Phi'(S_T) \frac{\partial S_T}{\partial x}]
\]

(13)

\[
e^{-rT} \mathbb{E}[\Phi(S_T) \frac{\partial S_T}{\partial x} \delta(u)]
\]

(14)

where \(\delta(u)\) is the Skorohod integral and \(u\) is a suitable process.

4. **Conclusion.** We are working on pricing by decomposition [11], [1].

References

to option pricing approximation, Finance and Stochastics, 16, 3, 03–422.

Economy, 81, 637–654.

[3] Choi, Sun-Yong, Fouque, Jean-Pierre, Kim, Jeong-Hoon, 2013, Option pricing under hybrid stochas-
tic and local volatility, Quantitative Finance, 13, 8, 1157–1165.

Elasticity of Diffusions, Unpublished Draft, Stanford University.


the Malliavin calculus, Finance and Stochastics, 8, 2, 161.

for stochastic volatility models, Quantitative Finance, 10, 2, 177–194.


[9] Heston, Steven L., 1993, A Closed-Form Solution for Options with Stochastic Volatility with Ap-
plications to Bond and Currency Options, Review of Financial Studies.

[10] Petrou, Evangelia and others, 2008, Malliavin Calculus in Lévy spaces and Applications to Finance,

Investment strategies with long memory: log-optimal solution and approximations by Stochastic Gradient

ZSOLT NIKÁ
EMAIL: nika.zsolt@itk.ppke.hu

Abstract. One major question in portfolio optimization is how to construct an optimal investment strategy when the stock price has long memory. In my presentation, I will present a family of models in discrete time, called Conditionally Gaussian model family, which can incorporate long memory such that log-optimal solution exists, and it can be calculated numerically.

Next, I will introduce two approximative solutions which are computationally feasible. The numerical solution in these cases can be obtained by simply evaluating a function instead of numerical integrals.

One of the approximative solutions suggests threshold type strategies. The essence of this strategy class is that, in this case, we are able to construct a stochastic gradient algorithm in such a way that it converges to the optimal strategy. The method works even if the stock price has long memory.

For numerical results, I will use a variant of the Fractional Stochastic Volatility model to demonstrate log-optimal solution. It is constructed in the spirit of the Fractional Stochastic Volatility model. This model has long memory, belongs to the Conditionally Gaussian family, and it has several statistical advantages. For showing results when we are using Stochastic Gradient method we rely on somewhat simpler processes.
Numerical Approximation of Dynamic Initial Margin: Chebyshev on the Market Space analysed

Maria R Nogueiras  
Traded Risk Analytics, HSBC  
Email: maria.r.nogueiras@hsbc.com

Ignacio Ruiz  
MoCaX Intelligence  
Email: i.ruiz@iruiztechnologies.com  
URL: http://https://www.mocaxintelligence.com/

Mariano Zeron  
MoCaX Intelligence  
Email: m.zeron@iruiztechnologies.com

Abstract.
Under BCBS-261 regulation bilateral counterparties (above a certain threshold volume) are required to exchange daily Initial Margin (IM) on a gross basis and in segregated accounts. This IM is intended to offset the gap risk still remaining after full collateralisation (daily exchange of Variation margin (VM) with zero thresholds). As a risk mitigant, IM increases the margin (defaulter paying) and decreases the capital (survival paying).

The industry adopted the ISDA proposed model (SIMM) for IM computation: sensitivity based, 99 percentile, 10 day VaR calibrated to historical data including at least 25% of stress data.

In the context of counterparty credit risk (CCR) computations, the existence of bilateral initial margin implies:
- Incremental funding cost, as a result of the posted IM, and the Margin Value Adjustment (MVA) as a new XVA

$$MVA(t) = \mathbb{E} \left( \int_t^T r_f(s)DF(t,s)IM(t)_{\text{posted}}(s)ds \right)$$

where \(r_f\) is the funding rate of the bank, \(DF\) is the discount factor
- Reduced exposures and therefore lower CVA (respectively, DVA) as a result of received IM (respectively, posted IM) and less CCR capital.

Assuming a single netting set covered by \(VM\) and \(IM\), banks B’s exposure to counterparty C can generally be represented as

$$E(t) = [V(t) - VM(t) - IM(t)_{\text{received}}]$$

with \(V\) the netting set value

The estimation of the above constitutes a computational burden as we need to compute a VaR type measure within the CCR Monte Carlo simulation or, more specifically, trade level sensitivities for every simulation path and time.

We will be reviewing the alternatives considered in the literature to address the problem, from approximations as scaled VaR or regression based techniques [3], [4], to brute force approach using adjoint differentiation [1].

From the different numerical approximation methods, Chebyshev on the Market Space [5] appears to offer an ideal balance between flexibility, accuracy and computational burden. In this paper we test such claim and we further explore the technique, extending existing reach to multi-factor diffusion models and a variety of financial products.
References


Abstract. Once the credit scoring, assigned by a financial institution to a client who enjoys a personal credit, is known, it is interesting to find the probability that the borrower declares himself unable to face the debt contracted with the bank after some time (for example, one year) of its granting. The main aim of this work is to propose models to estimate this probability, known as the probability of default and denoted by $PD$.

For this purpose, the variable “time to default” will be studied, taking advantage of the information that the credit scoring provides about it. However, throughout the study of a set of credits, the default is not observed for all of them: some credits fall into default during the study, but some do not and others may be cancelled or paid in advance. So, the variable “time to default” is censored.

On the other hand, the probability of default conditional to the credit scoring can be written as a transformation of the conditional survival function of the variable “time to default” by means of the following expression:

$$PD(t|x) = 1 - \frac{S(t + b|x)}{S(t|x)}$$

where $S(t|x)$ is the conditional survival function of time to default variable conditional to the credit scoring and $b$ is the horizon of default.

As a consequence of both these features, censored data techniques and survival analysis can be used in this context and they provide estimators of the conditional survival function that derive in estimators of the probability of default through the previous transformation.

Three estimators for the conditional survival function have been first considered: Beran’s generalized product-limit estimator ([1]), weighted local linear (WLL) estimator ([2]) and Van Keilegom-Akritas estimator ([3]). They have been transformed to obtain the corresponding $PD$ estimators.

A problem in the WLL estimator was detected. It is not designed to estimate a survival function that has to be non-increasing and whose range lies in the interval $[0, 1]$, so, using it in its original version, $PD$ estimations that take values greater than one or negative are obtained. As an alternative to this estimator, a local constant estimator with Nadaraya-Watson weights (WNW) is proposed. This estimator respects the isotony and the belonging to the interval $[0, 1]$ of the conditional survival function, returning correct estimations of the $PD$.

Under certain conditions and assuming that expressions of bias and variance of any conditional survival function estimator are had, a theoretical result that has been proven allows to obtain asymptotic properties of bias and variance of the corresponding $PD$ estimator. This is the case of the WLL and Van Keilegom-Akritas survival estimators for which asymptotic expressions of bias and variance are known ([2], [3]), thus it was possible to get results for the WLL and Van Keilegom-Akritas estimators of the probability of default. The asymptotic expressions of bias and variance for the Beran estimator of the $PD$ are known and were proven along with the asymptotic normality of the estimator in [4]. Nevertheless, these are complex expressions that make it difficult to obtain an approximation of the MSE or MISE since they depend on several unknown parameters.

In order to analyse the behaviour of the four $PD$ estimators by simulation, several models have been considered. In the first studies carried out, Beran estimator provides the best results because it gives the best approximations (smaller mean squared error) and less computation time. The WNW estimator
estimator reduces the mean squared error committed by WLL. The WLL and Van Keilegom-Akritas estimators for the distribution function turn out to be better than Beran’s with heavy censorship. A simulation study was carried out to check whether this property is inherited by the corresponding PD estimators. The results show that the Van Keilegom-Akritas estimator produces a smaller mean squared error in the right tail of the “time to default” distribution when the censorship is heavy. These four estimators involve a smoothing parameter in the covariate (credit scoring); the choice of this bandwidth parameter is also being examined in the simulation study.

Another fact observed in the simulation is that the PD estimators conditional to the credit scoring present excessive variability even using a large bandwidth to smooth the covariate. Therefore, a modification of the proposed estimators based on smoothing in the time variable is proposed. The first analyses seem to indicate that these estimators have a better behaviour.

References


Instantaneous volatility seasonality of high-frequency markets in directional-change intrinsic time

Vladimir Petrov  
University of Zurich, Switzerland  
Email: vladimir.petrov@uzh.ch

Anton Golub  
Flov Technologies AG

Richard Olsen  
Lykke Corp.

Abstract. We propose a novel intraday instantaneous volatility measure which utilises sequences of drawdowns and drawups non-equidistantly spaced in physical time as indicators of high-frequency activity of financial markets. The sequences are re-expressed in terms of directional-change intrinsic time which ticks only when the price curve changes the direction of its trend by a given relative value. We employ the proposed measure to uncover market microstructure weekly volatility seasonality patterns of three Forex and one Bitcoin exchange rates as well as a stock market index. We demonstrate the long memory of instantaneous volatility computed in directional-change intrinsic time. The provided volatility estimation method can be adapted as a universal multiscale risk-management tool independent of the discreteness and the type of analysed large high-frequency data.

Keywords: instantaneous volatility; market microstructure; seasonality; high frequency markets; risk management; computational finance
Accelerated Share Repurchase: option hedging, optimal execution and neural networks

JIANG PU
INSTITUT EUROPLACE DE FINANCE, 28 PLACE DE LA BOURSE, 75002 PARIS, FRANCE
EMAIL: jiang.pu.2009@m4x.org

Abstract. When firms want to repurchase their own shares, they often enter an Accelerated Share Repurchase (ASR) contract with a bank. The bank buys the shares for the firm and is paid the average market price over the execution period, the length of the period being decided upon by the bank during the buying process. Mathematically, the optimization problem faced by the bank consists of both option hedging (Asian and Bermudan options) and optimal execution. In our previous work (see [1]), we solved the Bellman equations associated with the problem with a change of variables related to Asian options that reduced the dimension of the problem. Other variants of ASR exist and new features appeared in the recent contracts. The optimization problem is more complex and cannot be handled using the same dimension reduction technique. We propose to use artificial neural networks to tackle the problem of dimensionality.

1. ASR contract. For the sake of simplicity of the presentation, we focus here on the ASR with fixed number of shares. An ASR with fixed number of shares $Q$ and with maturity $T$ is the following contract written between a firm and a bank:

- At time $t = 0$, the bank borrows $Q$ shares from shareholders (usually institutions) and gives the shares to the firm in exchange of a fixed amount $QS_0$ where $S_0$ is the MtM price of the stock at time $t = 0$. The bank then has to progressively buy back $Q$ shares on the market to give them back to the lenders, thus, goes from a short position to a flat position on the stock.

- The bank is long an option with payoff $Q(A_\tau - S_0)$ (the firm being short of this option), where $A_t$ denotes the average price between 0 and $t$ (in practice the average of closing prices or the average of daily VWAPs), and where $\tau$ is a stopping time chosen by the bank among a set of contractually specified dates.

Hence, the firm eventually pays the price $A_\tau$ for each of the $Q$ shares that have been repurchased. Although from the banks perspective, entering an ASR contract can be considered being long a Bermudan option, classical derivative pricing theory is not suited to solve the problem of hedging such a contract. Indeed, on the one hand, the option hedging problem is coupled with an optimal execution problem; and on the other hand, liquidity plays a crucial role here.

2. Optimization Problem. We consider a discrete model where each period of time corresponds to one day (of length $\delta t$). In other words, if the interval $[0, T]$ corresponds to $N$ days ($T = N\delta t$), we consider the subdivision $(t_n)_{0 \leq n \leq N}$ with $t_n = n\delta t$ and we index variables with $n$.
We define:

- the market daily VWAP $S_n$ at day $n$; we do not need to specify the dynamics of $(S_n)$ at this stage;
- the running average $A_n = \frac{1}{n} \sum_{k=1}^{n} S_k$, $n \geq 1$;
- the filtration $(\mathcal{F}_n)$ generated by $(S_n)$;

\[1\]See [2] for instance
• the amount of shares bought \( v_n \) on day \( n \), \( n \geq 1 \);
• the cumulative number of shares bought \( q_n = \sum_{k=1}^{n-1} v_k \);
• the cash account \( X_n = -\sum_{k=1}^{n-1} (v_k S_{k+1} + L (\frac{F}{L}) V) \) where \( L \) is the execution cost function and \( V \) the market average daily volume;
• the set of possible dates of early exercise (maturity not included) \( \mathcal{N} \).

An admissible strategy is determined by:

• the amount to be bought at each day, \( \theta(n) \);
• a stochastic stopping policy \( (p_n) \), which is an \( \mathcal{F}_n \)-adapted process that takes values in \([0,1]\), \( n \in \mathcal{N} \). The effective stopping is denoted by \( \pi_n \in \{0,1\} \), it is conditionally independent of \( \pi_n' \) for \( n' < n \).

The PnL of the strategy is then given by:

\[
PnL = \sum_{n=1}^{N} \prod_{k=1}^{n-1} (1 - \pi_k) \pi_n (X_n + QA_n + (q_n - Q)S_n) + \ell(q_n - Q),
\]

where \( \ell \) is a penalization function.

We search for the optimal strategy \( v \) in the form of \( v_n = v^\theta(n, S_n, A_n, q_n) \), and \( p \) in the form of \( p_n = p^\phi(n, S_n, A_n, q_n) \) for \( n \in \mathcal{N} \), both of them lie in a finite dimensional space of functions parameterized by \( \theta \) and \( \phi \) respectively.

In practice, we will use neural networks with some normalizations adapted to the problem.

The problem is to optimize a performance measure such as the mean-variance criterion\(^2\):

\[
\max_{\theta, \phi} MV = \max_{\theta, \phi} \left\{ E[PnL] - \frac{\gamma}{2} V[PnL] \right\}. \]

In practice, trajectories of \( (S_n) \) will be sampled. The performance measure can be approximated by its Monte Carlo estimator:

\[
MV (\theta, \phi; (S')) = \frac{1}{T} \sum_{i=1}^{T} \left[ \frac{1}{N} \sum_{n=1}^{N-1} \prod_{k=1}^{n-1} (1 - p^i_k) p^i_n PnL^i_n \right. \\
- \frac{\gamma}{2} \left( \frac{1}{T} \sum_{i=1}^{T} \sum_{n=1}^{N-1} \prod_{k=1}^{n-1} (1 - p^i_k) p^i_n (PnL^i_n)^2 - \left( \frac{1}{T} \sum_{i=1}^{T} \sum_{n=1}^{N-1} \prod_{k=1}^{n-1} (1 - p^i_k) p^i_n PnL^i_n \right)^2 \right),
\]

where \( (S_n)^i \) denotes the \( i \)th trajectory of \( (S_n) \) sampled, and for every other \( \mathcal{F}_n \)-measurable variable \( Y \) defined above, \( Y^i \) denotes its realization given \( (S_n) = (S_n)^i \).

The above criterion can be differentiated with respect to \( (\theta, \phi) \) using automatic differentiation tool, which enable us to optimize over the parameters.

Note that \( v^i_n = v^\theta(n, S_n, A_n^i, q_n^i) \) depends on \( \theta \) explicitly through the parameter of \( v^\theta \), but also implicitly through the fourth argument (the same applies to \( p^i_n \)), the network is said to be fully-recurrent.

\(^2\)Other risk measures could be used, for instance CVaR or expected utility.
3. Performance. We tested the method described above using the following setting:

- \( S_0 = 45, Q = 20\,000\,000; \)
- \( N = 63, N = \{22, \cdots, 62\}; \)
- \( V = 5\,000\,000, L : x \mapsto 0.1 \cdot x^{1.75}, \ell : x \mapsto 2 \cdot 10^{-7} \cdot x^2; \)
- \((S_n)^i\) are generated using i.i.d innovations of law \( \mathcal{N}(0, \sigma^2) \) with \( \sigma = 0.6 \) (which correspond to approximately 21% of annual volatility);
- \( v_\theta \) is specified as:
  \[
v_\theta(n, S, A, q) = \left(1 + \tilde{v}_\theta\left(\frac{n}{N} - \frac{1}{2}, \frac{S - S_0}{S_0}, \frac{A - S_0}{S_0}, \frac{q}{Q} - \frac{1}{2}\right)\right) \cdot Q \frac{n}{N} - q,
\]
  where \( \tilde{v}_\theta \) is a neural network consisting of 4 inputs, a hidden layer of 50 neurons with ReLU activator and 1 output;
- \( p_\phi \) is specified as:
  \[
p_\phi(n, S, A, q) = S\left(\nu_\phi \cdot \left(\frac{q}{Q} - \tilde{p}_\phi\left(\frac{n}{N} - \frac{1}{2}, \frac{S - S_0}{S_0}, \frac{A - S_0}{S_0}, \frac{q}{Q} - \frac{1}{2}\right)\right)\right) \cdot 1_{n \in \mathbb{N}} + 1_{n = N},
\]
  where \( \tilde{p}_\phi \) is a neural network consisting of 3 inputs, a hidden layer of 50 neurons with ReLU activator and 1 output, \( \nu_\phi \) is a scalar parameter, and \( S \) the function defined by:
  \[
  S : x \mapsto \min\left(\max\left(\frac{2}{1 + e^{-x}} - \frac{1}{2}, 0\right), 1\right).
  \]

The learning curve is presented below. Each step consists of a batch of 1000 trajectories.

![Learning curves](image)

Figure 1: Learning curve, the y-axis corresponds to the mean-variance criterion normalized by \( QS_0 \)

Note that for large values of \( \gamma \), one needs to pre-train the networks with \( \gamma \) set to 0, and small execution costs and penalties, otherwise only the local optimum is reached, which corresponds to the trivial strategy consisting of buying the same number of shares every day and deliver them at expiry.

References

International Journal of Theoretical and Applied Finance, 18(03), 1550019.
Journal of Risk, 19(5).
Robo-advising: unfolding the risks

PAULO M. SILVA
ISEG, UNIVERSIDADE DE LISBOA, PORTUGAL
LUSITANIA SEGUROS, S.A., PORTUGAL
EMAIL: paulomartinssilva@phd.iseg.ulisboa.pt, paulo.martins.silva@lusitania.pt

Abstract. An important aspect of robo-advising is the level of behavioural elements associated with an investment decision making process. Most automatic platforms agree that understanding investor psychology and investors’ expectations is important to create a sound financial strategy [4].

A relevant argument is the distinction between clients’ decisions, which are driven by their preferences, and those that are driven by psychological biases. This is also the reason why it is important for advisors to understand how their clients make financial decisions. Not understanding behaviour biases may result in inappropriate financial advising.

This perspective is important for establishing a link between the clients’ needs and certain financial objectives because the typical perception is that traditional finance defines a rational benchmark, and all behaviourally motivated decisions of the clients are irrational. This is the frontier between a framework for rational and irrational investors. Advising is helping clients to take rational decisions, which are consistent with their needs and preferences. Although some theories define a strict framework for rationality, this may not fit clients’ needs [10, 11, 12, 13].

In the industry the challenge is how advisors calibrate the theoretical framework with a sound and clear classification on clients’ needs and perceptions [5]. Another frontier is how algorithms used by robot-advising incorporate these features on their questionnaires. In fact, there are clear distinctions between the results from the robo-advisors, either from a client risk profiling or from an asset allocation perspective. These outcomes are a major concern for individual investors that have no financial literacy to perform back testing of each proposed strategy by the robot. In spite of individual investors being alerted either by financial market regulators or trading regulators (e.g., [3], SEC, PRA, ...), and regardless the fact that some markets, as the European, being under the umbrella of strict directives (i.e., MIFID), there are relevant risks of mismatching between investors goals and solutions provided by robo-advisors, due to some large scope of understanding within the strategies put forward by the algorithms [7].

Over the last several years, with an ageing population transferring assets to millennial, individuals are more susceptible to utilizing internet-based platforms for most of their financial needs. Large investment institutions are mindful of these trends and have started investing in the infrastructure of robo-platforms [6]. As artificial intelligence evolves, the software will be designed to handle more complex, tougher strategies, whereas most of the current robo-platforms are not equipped to handle these scenarios. The strategies are offered based on Exchange Tradable Fund (ETF), and there are questions on how based investing performs would survive during a financial crisis as it would present itself the next time. Nonetheless, the levels of trust humans are willing to have on a fully automated investment platform will determine how fast robo-advising is integrated, especially on the event of financial crises and the way conflicts of interests are managed [7, 1].

The robo-advisor platforms consider factors that are significant to an investor’s financial needs. These recommendations are completely determined by the depth of questionnaires used by each platform. A second factor is how thorough the answers to the questions are and if the pre-selected fields used to answer the questions are 100% matching the client’s individual situation. The roboplatforms also consider the type of account an investor is seeking advice for.

However, in relation to investors’ protection, either on the capability of the financial services firms providing robo-advising, or at the wealth investment security, there are limitations that individual investors should be aware of [3]. These are related with safety of investments and with the necessary due diligences on financial firms. Although financial literacy is a relevant topic on financial decisions, there is potential for the inability to correctly describe the relationship between investors and robo-advisor firms [9] specially on fiduciary duties. The potential to better alignment between investor targets and robo-advisors is being tested, as its popularity is growing, but there are signs that portfolio management improvements are possible and some of the pitfalls derived from pervasive
behavioral biases can be reduced [2, 8].

Robo-advising platforms although still within their early phases of development, have gained quite a bit of traction within the wealth management sector. From our analysis we expect robo-advisory platforms will continue to evolve over the years. As technology evolves, the robo-advisory platforms may continue to experience an up-tick. This could be affected given the millennial preference for online based services. If the focus remains on cost cutting elements, then passive indexing strategies could remain a viable option. Another important component is the regulatory landscape and its evolution. In this sense, the focus that regulatory frameworks take into the role of asset managers, depositories, custodians, exchanges, must go beyond the nature of the supervisory of markets to the conduct of the firms managing the robo-advisory platforms.

The value chain of these emergent financial services is unclear regarding the fiduciary duties and is necessary a clarification due to the velocity of technological evolution. However, risk profiling is still a very grey area when financial literacy is not robust enough for investors to comprehend the pitfalls of a wider class of protective strategies. The existence of stylized investors with different risk profiles allowing the industrialization of the asset allocation process, may limit the capital protection strategies for a large number of investors. However, we are seeing innovation using artificial intelligence and machine learning that can trigger the offer of downside protective strategies. These areas are evolving and, along with regulation, there is a need to deeper research on such dynamic topics especially if key risk indicators are to be included into investors decisions and these are not coherent risk metrics nor predictive of future volatility.

References


Efficient numerical valuation of high-dimensional basket options via partial differential equations

JACOB SNOEIJER*
MIDDELHEIMLAAN 1, B-2020 ANTWERP, BELGIUM
EMAIL: jacob.snoeijer@uantwerpen.be

KAREL J. IN ’T HOUT*
MIDDELHEIMLAAN 1, B-2020 ANTWERP, BELGIUM
EMAIL: karel.inthout@uantwerpen.be

Abstract. This talk deals with numerical methods to approximate the fair values of European and Bermudan basket options, which constitute common products in the financial markets. If there are \( d \geq 2 \) assets in the basket, then the fair value of such a financial option satisfies a time-dependent \( d \)-dimensional partial differential equation. For its efficient numerical solution, we combine the principal component analysis \([2]\) with modern alternating direction implicit (ADI) schemes \([1]\). The convergence of this approach is investigated in detail, and ample numerical experiments are presented illustrating the high performance it attains.

A European option is a contract that gives its holder the right but not the obligation to buy or sell an underlying asset, at a specified future date \( T \) for a specified strike price \( K \). The fair value of an option is derived from the price of the underlying asset (e.g. a stock or a foreign exchange rate). At maturity time \( T \) the value of the option is known, and given by the pay-off \( \phi(s) \).

Instead of an option on one asset, we are interested here in options on a basket of \( d \geq 2 \) assets. For example, a put option on a weighted average of \( d \) assets has pay-off function

\[
\phi(s_1, s_2, \ldots, s_d) = \max \left( K - \sum_{i=1}^{d} \omega_i s_i, 0 \right) .
\]

Under the Black-Scholes model, one arrives at a \( d \)-dimensional partial differential equation (PDE) for the fair option value \( u(s_1, s_2, \ldots, s_d, t) \):

\[
\frac{\partial u}{\partial t} = \frac{1}{2} \sum_{i=1}^{d} \sum_{j=1}^{d} \rho_{ij} \sigma_i \sigma_j s_i s_j \frac{\partial^2 u}{\partial s_i \partial s_j} + \sum_{i=1}^{d} r_i s_i \frac{\partial u}{\partial s_i} - ru
\]

\( u|_{t=0} = \phi(s_1, s_2, \ldots, s_d) \).

We will investigate a method by Reisinger & Wittum \([2]\) that efficiently approximates the solution to this \( d \)-dimensional PDE by a linear combination of solutions to lower-dimensional PDEs. This method will be applied to approximate the fair values of European and Bermudan basket options on many assets. Next, a detailed numerical study will be performed of the stability and convergence behaviour of this approach.

References


*Department of Mathematics and Computer Science, University of Antwerp
Multi-curve interest rate modeling with Hull-White and Black-Karasinski

Sergei Vassilovski
Quantitative Research Group, Risk Analytics, IBM
Email: sergei.vassilovski1@ibm.com

Alex Kreinin
Quantitative Research Group, Risk Analytics, IBM
Email: alex.kreinin@ca.ibm.com

Abstract. After the 2008 financial crisis, using a single interest rate curve for both discounting and pricing of LIBOR-based interest rate derivatives is no longer an option. A number of multi-curve frameworks have been proposed since then where the collateral (OIS) curve is used for discounting and different forwarding LIBOR curves are used for pricing interest rate swaps with different tenors (see, for instance, [1], [2], [3]).

Here we present a new multi-curve interest rate model with stochastic basis where the OIS curve is modeled using one factor Hull-White model, whereas the spot LIBOR-OIS spread is modeled using a regularized Black-Karasinski model. It is based on the defaultable bond model in [4]. One of the immediate advantages of using the regularized BK model for the LIBOR-OIS spread is that the resulting stochastic spread is always positive. Under this model we obtain semi-analytic formulas for forward LIBOR rates, as well as OIS-discounted prices of collateralized swaps, swaptions, and caplets. The pricing formulas can be easily computed numerically. It also enables us to derive the risk-neutral dynamics of forwarding LIBOR curves for different tenors without imposing any additional assumptions on their spreads.

1. Introduction. Let $R(t)$ denote the collateral (OIS) short rate, and let $Q$ be the risk-neutral measure under which the price $V(t)$ of any collateralized financial instrument grows at the rate $R(t)$, or, in other words, its discounted price, $\exp\left(-\int_0^T R(u)du\right) \cdot V(t)$, is a martingale under $Q$. We define the collateral discount curve $P(t,T)$ at time $t \geq 0$ as follows:

$$P(t,T) = \mathbb{E}_Q \left[ \exp\left(-\int_t^T R(u)du\right) \mid \mathcal{F}_t \right], \quad t \leq T.$$  

We also assume that, under measure $Q$, the collateral rate $R(t)$ follows the one-factor Hull-White model with constant volatility, i.e.,

$$R(t) = \varphi(t) + Y(t),$$  

where

$$dY(t) = -k_y Y(t)dt + \sigma_y dW^y(t), \quad Y(0) = 0,$$

and the deterministic function $\varphi(t)$ is chosen in a way that allows a perfect fit into the initial term structure of the collateral discount curve $P(0,T)$.

Let

$$B(t,T) = \mathbb{E}_Q \left[ \exp\left(-\int_t^T \lambda(u) + R(u)du\right) \mid \mathcal{F}_t \right], \quad t \leq T,$$

be the discount curve used in the definition of the the spot $\delta$-period LIBOR rate

$$L(t,t+\delta) = \frac{1}{\delta} \left( \frac{1}{B(t,t+\delta)} - 1 \right)$$.
for any tenor $\delta$ at time $t$. The forward LIBOR rate $L(t,T_1,T_2)$ at time $t$ for the future time period $[T_1,T_2]$ is defined as

$$L(t,T_1,T_2) = \frac{1}{P(t,T_2)} \cdot \mathbb{E}_Q \left[ \exp \left( - \int_t^{T_2} R(u) du \right) L(T_1,T_2) \bigg| \mathcal{F}_t \right], \quad t \leq T_1 < T_2.$$  

The LIBOR-OIS spread $\lambda(t)$ is modelled using a regularized Black-Karasinski model:

$$\lambda(t) = \varphi_\lambda(t) + \lambda_0 \cdot \exp(X(t) - \alpha X(t)^2),$$

where the factor $X(t)$ follows, under measure $Q$, a zero-level Ornstein-Uhlenbeck process with constant volatility

$$dX(t) = -k_x \cdot X(t) dt + \sigma_x dW_x(t), \quad X(0) = 0,$$

$\lambda_0 > 0$ is a constant, and the deterministic function $\varphi_\lambda(t)$ is chosen in a way that allows a perfect fit into the initial term structure of the LIBOR discount curve $B(0,T)$. The regularizer $-\alpha X(t)^2$, where $\alpha > 0$, is used to ensure convergence of the improper integrals that appear in the pricing formulas below. We denote by $\rho$ the correlation coefficient between $W^y(t)$ and $W^x(t)$:

$$dW^y(t) \cdot dW^x(t) = \rho \cdot dt.$$  

The key role in the approximate solution to the model plays the following deterministic function

$$\mathcal{E}(x,t,T) = \mathbb{E}_Q \left[ \exp \left( \left\{ - \int_t^T \lambda_0 \exp(X(u) - \alpha X(u)^2) du \right\} 
\right. 
\left. + \left\{ - \int_t^T \rho \bar{Y}(u) du - \frac{1}{2} \text{Var} \left( - \int_t^T \rho \bar{Y}(u) du \right) \right\} \bigg| X(t) = x \right],$$

with $X(t) = x$ and

$$d\bar{Y}(u) = -k_y \bar{Y}(u) du + \sigma_y dW_y(u), \quad \bar{Y}(t) = 0, \quad u \geq t,$$

where $W^y(u)$ is the Brownian motion process driving $X(u)$. There is no known closed-form expression for the expectation, but there are quite accurate semi-closed approximation formula(s) for the function $\mathcal{E}(x,t,T)$ in [5]. Let

$$BE(t,T) = \frac{B(0,T)}{B(0,t)} \cdot \frac{\mathcal{E}(0,0,t)}{\mathcal{E}(0,0,T)}$$

and

$$I(t,T) = \frac{P(0,t)}{P(0,T)} \cdot BE(t,T) = \frac{P(0,t)}{P(0,T)} \cdot \frac{B(0,T)}{B(0,t)} \cdot \frac{\mathcal{E}(0,0,t)}{\mathcal{E}(0,0,T)}.$$  

Also let, for $k > 0$ and $\tau \geq 0$,

$$H(k,\tau) = \frac{1 - \exp(-k\tau)}{k}.$$  

Then the values of the discount curves $P(t,T)$ and $B(t,T)$ depend only on the values of the stochastic processes $X(t)$ and $Y(t)$ at time $t \geq 0$:

$$P(t,T) = P(Y(t),t,T) \quad \text{and} \quad B(t,T) = B(X(t),Y(t),t,T),$$

where

$$P(y,t,T) = \frac{P(0,T)}{P(0,t)} \cdot \exp \left( -H(k_y,T-t) \cdot \left\{ y + \frac{1}{2} \sigma_y^2 \cdot [H(k_y,t)^2 + H(k_y,T-t) \cdot H(2k_y,t)] \right\} \right)$$

and

$$B(x,y,t,T) = I(t,T) \cdot P(y,t,T) \cdot \mathcal{E}(x,t,T).$$
We will also need the functions
\[cx(x,z,\tau) = e^{-k_x \tau} x - \rho \sigma_x \sigma_y \frac{H(k_x, \tau) - H(k_x + k_y, \tau)}{k_y} + \sigma_x \sqrt{H(2k_x, \tau)} \cdot z,\]
\[cx(x,z_1,z_2,\tau) = e^{-k_x \tau} x - \rho \sigma_x \sigma_y \frac{H(k_x, \tau) - H(k_x + k_y, \tau)}{k_y} + \sigma_x \left( \rho \frac{H(k_x + k_y, \tau)}{\sqrt{H(2k_y, \tau)}} \cdot z_1 + \sqrt{H(2k_y, \tau)} - \rho^2 \frac{H(k_x + k_y, \tau)^2}{H(2k_y, \tau)} \cdot z_2 \right),\]
\[cy(y,z,\tau) = e^{-k_y \tau} y - \sigma_y \frac{H(k_y, \tau) - H(2k_y, \tau)}{k_y} + \sigma_y \sqrt{H(2k_y, \tau)} \cdot z.\]

2. Pricing Formulas.

2.1. Forward LIBOR Rates.

Theorem 2.1. For \(0 \leq t \leq T\), the forward LIBOR rate \(L(t, T, T + \delta)\) satisfies
\[L(t, T, T + \delta) = L(X(t), Y(t), t, T, T + \delta),\]
where the deterministic function \(L(x, y, t, T_1, T_2)\) is defined by
\[L(x, y, t, T_1, T_2) = \frac{1}{T_2 - T_1} \cdot \left( \frac{P(y, t, T_1)}{P(y, t, T_2)} \cdot \frac{1}{I(T_1, T_2)} \cdot \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp(-z^2/2) dz \right) - 1.\]

2.2. Interest Rate Swap.

Theorem 2.2. Consider an interest rate swap whose floating leg pays at time \(T_k\) the LIBOR rate set at time \(T_{k-1}\) with tenor \(\Delta T_k = T_k - T_{k-1}\):
\[\Delta T_k \cdot L(T_{k-1}, T_k), \quad k = 1, 2, \ldots, n.\]

And let the fixed leg of the swap pays a fixed rate \(K\) on dates \(\hat{T}_i\), \(i = 1, 2, \ldots, m\). Then the value of the swap to the fixed rate payer equals
\[IRS(t, K) = IRS(X(t), Y(t), t, K),\]
where the deterministic function \(IRS(x, y, t, K)\) is given by
\[IRS(x, y, t, K) = \sum_{k=1}^{n} \Delta T_k \cdot P(y, t, T_k) \cdot L(x, y, t, T_{k-1}, T_k) - K \sum_{i=1}^{m} \Delta \hat{T}_i \cdot P(y, t, \hat{T}_i),\]
with the function \(L(x, y, t, T_1, T_2)\) being defined in Theorem 2.1.

2.3. Payer Swaption.

Let the deterministic function \(IRS(x, y, z_1, z_2, t, T)\) be given by
\[IRS(x, y, z_1, z_2, t, T) = IRS(cx(x, z_1, z_2, t-T), cy(y, z_1, T-t), T, K),\]
where the function \(IRS(x, y, t, K)\) is defined in Theorem 2.2. Then the following theorem holds

Theorem 2.3. The price at time \(t\) of a payer swaption that gives its holder the right to enter at time \(T_0 \geq t\) into an interest rate swap whose floating leg pays at time \(T_k\) the LIBOR rate set at time \(T_{k-1}\) with tenor \(\Delta T_k = T_k - T_{k-1}\):
\[\Delta T_k \cdot L(T_{k-1}, T_k), \quad k = 1, 2, \ldots, n,\]
and the fixed leg pays the fixed rate \(K\) on dates \(\hat{T}_i\), \(i = 1, 2, \ldots, m\), equals
\[OIRS(t, K) = \mathbb{E}_Q \left[ \exp \left( - \int_{T_0}^{T} R(u) du \right) \cdot IRS(T_0, K) \right] \bigg| \mathcal{F}_t \]
\[= P(Y(t), t, T_0) \cdot \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} IRS(X(t), Y(t), z_1, z_2, t, T_0, K)^+ \cdot \exp \left( - \frac{z_1^2 + z_2^2}{2} \right) dz_1 dz_2.\]
2.4. **Caplet.** Let, for $0 \leq t \leq T_1 < T_2$,

$$\Delta T = T_2 - T_1.$$ 

$$\bar{\sigma}(t, T_1, T_2) = \sigma_y \cdot H(k_y, \Delta T) \cdot \sqrt{H(2k_y, T_1 - t) - \rho^2 \frac{H(k_x + k_y, T_1 - t)^2}{H(2k_x, T_1 - t)}}.$$ 

$$K(x, z, t, T_1, T_2, k) = \frac{1}{(1 + k\Delta T) \cdot BE(T_1, T_2) \cdot \hat{e}(\mathbf{e}(x, z, T_1 - t), T_1, T_2)}.$$ 

$$F(y, z, t, T_1, T_2) = \exp\left(\frac{\bar{\sigma}(t, T_1, T_2)^2}{2} - H(k_y, \Delta T) \cdot \left[ e^{-k_y(t - T_1)} y - \sigma_y^2 H(k_y, T_1 - t) - H(2k_y, T_1 - t) \right] \right),$$

$$+ \sigma_y \cdot \rho \cdot \frac{H(k_x + k_y, T_1 - t)}{\sqrt{H(2k_x, T_1 - t)}} \cdot \left[ z + \frac{1}{2} \sigma_y^2 \cdot (H(k_y, T_1)^2 + H(k_y, \Delta T) \cdot H(2k_y, T_1)) \right],$$

$$d_-(x, y, z, t, T_1, T_2, k) = \log(F(y, z, t, T_1, T_2)/K(x, z, t, T_1, T_2, k)) - \sigma(t, T_1, T_2)^2/2,$$

$$d_+(x, y, z, t, T_1, T_2, k) = \log(K(x, z, t, T_1, T_2, k)/F(y, z, t, T_1, T_2)) + \sigma(t, T_1, T_2)^2/2,$$

$$G(x, y, z, t, T_1, T_2, k) = K(x, z, t, T_1, T_2, k) \cdot \Phi(-d_-(x, y, z, t, T_1, T_2, k)) - F(y, z, t, T_1, T_2) \cdot \Phi(-d_+(x, y, z, t, T_1, T_2, k)),$$

where

$$\Phi(a) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{a} \exp(-u^2/2)du, \ a \in (-\infty, \infty).$$

Then we obtain

**Theorem 2.4.** The price $c_t$ at time $t$ of a caplet with a reset date $T_1 \geq t$ and a settlement date $T_2 > T_1$ that, at time $T_2$, pays the holder the positive part of the difference between the simple spot rate $L(T_1, T_2)$ and the strike $k$: 

$$\Delta T \cdot (L(T_1, T_2) - k)^+, \text{ where } \Delta T = T_2 - T_1,$$

equals

$$c_t = E_Q \left[ \exp\left(-\int_{T_1}^{T_2} R(u)du\right) \cdot \Delta T \cdot (L(T_1, T_2) - K)^+ \bigg| F_t \right]$$

$$= (1 + k\Delta T) \cdot P(t, T_1) \cdot \frac{P(0, T_2)}{P(0, T_1)} \cdot \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} G(X(t), Y(t), z, t, T_1, T_2, k) \exp(-z^2/2)dz.$$ 

**References**


220
Parameter estimation of affine term structure models

SAMYUKTA VENKATARAMANAN
ROOM KA1-009 ANNEXE BUILDING, KEVIN STREET, TECHNOLOGICAL UNIVERSITY DUBLIN, IRELAND
EMAIL: d14124600@mydit.ie

STEPHEN O’SULLIVAN
ROOM KA1-013, ANNEXE BUILDING, KEVIN STREET, TECHNOLOGICAL UNIVERSITY DUBLIN, IRELAND
EMAIL: stephen.osullivan@dit.ie

Abstract. The econometric analysis of continuous-time affine asset pricing models requires the fast and accurate estimation of parameters. In parameter estimation, the density functions of the discretely sampled returns are obtained by solving a set of partial differential equations, which is computationally challenging as we need to compute the density function repeatedly on a large domain.

Feuerverger (1990) [1] proposed the basic idea of estimating parameters for a given stationary time series using the Characteristic Function (CF). Building on this, Singleton (2000) [8] proposed a framework for parameter estimation using the CF of a stochastic process. The CCF for a discretely sampled time series from an $N$-dimensional affine diffusion, denoted as $\phi(u, \gamma)$, is dependent on the scalar transform variable $u$ and vector of model parameters $\gamma$. The proposed method utilizes $E[e^{iu\cdot Y_{t+1}}|Y_t] = \phi(u, \gamma)$, to develop a method-of-moments estimator directly in the frequency domain, negating any need for computationally expensive Fourier inversions. This method of exploiting the Empirical CCF (ECCF) is more efficient than the traditional maximum-likelihood estimators for problems where the closed form of the conditional densities of discretely sampled returns are not known.

The functional form of the CF for affine processes have nearly closed-form expressions. That is, we can obtain compute it by solving a system of differential equations, namely the so-called Riccati equations which are stiff. Solving this set of stiff differential equations repeatedly for each iteration in our parameter estimation builds up a considerable computational cost. In general, implicit solvers are preferred over explicit solvers for solving stiff systems. This is mainly because implicit methods surpass the explicit in terms of efficiency, despite the computational complexity of matrix inversion. Recently, for mildly stiff problems a class of Extended Stability Runge-Kutta (ESRK) methods have been introduced which have considerably less complexity than their implicit counterparts, but still maintaining superior efficiency over explicit techniques. One such class is the Runge-Kutta-Gegenbauer (RKG) stability polynomials proposed by O’Sullivan (2017) [7]. RKG stability polynomials have been shown to generate stability domains with imaginary extent which is crucial for non-symmetric problems. Hence, they are an extension on SRK, which have their extended stability domain on the negative real axis, and provide greater stability properties. The polynomials have a arbitrarily high order of accuracy and the range of the stability domain in the imaginary direction is determined by the Gegenbauer parameter $\nu$. We present our findings of applying these stabilized explicit solvers for our parameter estimation problem and compare the efficiency with the traditional implicit and explicit solvers.

References


Interconnectedness of spot electricity prices. A dynamic network analysis

Guan Yan
Dept. of Actuarial Studies and Business Analytics, Macquarie University, Australia
Email: guan.yan@hdr.mq.edu.au

Stefan Trück
Dept. of Actuarial Studies and Business Analytics, Macquarie University, Australia
Email: stefan.trueck@mq.edu.au

Abstract. This study constructs dynamic Granger causality networks based on the spot electricity prices of five regional markets in the Australian National Electricity Market (NEM). Based on a data set comprising electricity spot prices from 1 July 2010 to 30 June 2017, we employ principal components analysis and generate Granger causality networks to examine the degree of interconnectedness of the NEM in a time-varying setting. We find that the derived measures of interdependence can be related to actual market events such as price spikes, unexpected high demand for electricity, sudden increase in price volatility, rebidding of dominant generators, the temporary of permanent outage of major power stations, as well as upgrades and limitations in transmission capacity. In the analysed network, we find that the stronger dependence is exhibited by regional markets that are linked by interconnectors, while the direction of Granger causality can be related to interregional trade. Furthermore, this study examines the usefulness of the derived measures as early-warning indicators for upcoming periods of extreme prices and volatility. Overall, our results suggest limited predictive power of the interconnectedness measures for the spot price behaviour in the NEM.

Keywords: Electricity markets, spot prices, dynamic networks, multivariate analysis, dependence and interconnectedness, risk management
Path-dependent leveraged exchange-traded fund option pricing

WENJUN ZHANG
DEP. OF MATHEMATICAL SCIENCE, AUCKLAND UNIVERSITY OF TECHNOLOGY, NEW ZEALAND
EMAIL: wzhang@aut.ac.nz

JILING CAO
GAURAV KAPOOR
XINFENG RUAN

Abstract. The growth of leveraged exchange-traded fund (ETF) industry has given rise to the trading of options written on leveraged ETFs. Catering to the industry needs, we consider the problem of pricing European options (which essentially are path-dependent) on leveraged ETFs. After introducing a new path-dependent process, we obtain a closedform formula for a European call option on a leveraged ETF. Our pricing formula is modified according to Carr and Madan (1999), and it can be implemented with the fast Fourier transform (FFT) approach in a friendly way. Our numerical analysis shows that the FFT approach is more accurate and efficient than the Monte Carlo (MC) method, in calculating the prices of European call options on leveraged ETFs.

References

Index

Abbas-Turki, 126
Abramova, 93
Agarwal, 125
Aïd, 23
Albani, 77
Alonso-García, 131
Álvarez, 135
Arratia, 135, 139
Arregui, 107

Bayer, 143
Ben Hammouda, 143
Boën, 97
Bokanowski, 102
Bondarenko, 144
Borovykh, 146
Boudabsa, 62
Bunn, 93

Calvo-Garrido, 89
Cao, J., 224
Cao, R., 208
Carlini, 58
Cerný, 114
Chen, 98
Cheng, 147, 149
Christara, 98
Cirillo, 147, 149, 155
Company, 115
Coulon, 91
Crépey, 24, 79
Cruz, 111, 113

de Frutos, 154
Debrabant, 102
Deelstra, 25
Deng, 124

Dorador, 139
Dupire, 26, 78
Düring, 103

Egorova, 115
Ehrhardt, 89, 128
El-Khatib, 201
Eliazar, 155
Elie, 24
Esquível, 39

Filipović, 59, 62
Fontanari, 155
Forde, 60
Forsyth, 27, 179
Fries, 64

Gaspar, 157
Gatón, 154
Gerhold, 60
Glau, 55, 63
Gobet, 121, 125
Golub, 210
Gómez, 70
Gonon, 84
Graf von Luckner, 30
Grzelak, 161
Gudkov, 169
Guerra, J., 43
Guerra, M., 43
Gulisashvili, 165
Guyon, 28, 80

Huré, 123
Hüser, 173

Iglesias, 178
in ’t Hout, 97, 216
Insley, 179
Jeanblanc, 29
Jevtic, 67
Jódar, 115
Kapllani, 128
Kapoor, 224
Khalilzadeh, 59
Kiesel, 30
Kikuchi, 180
Kilianová, 112
Kingston, 144
Kirkby, 195
Kramer, 30
Krasii, 39
Kreinin, 217
Kressner, 55
Kwak, 67
Lagunas, 165
Le Floch, 184
Lee, G., 190
Lee, S., 194
Lee, Y., 194
Leitao, 195
Li, L., 29
Li, X., 124
Lionnet, 127
Liu, 200
Londoño, 70, 74
Lopes, 44
López-Salas, 121
Makumbe, 201
Marchenko, 85
Melicherčík, 114
Merino, 165
Monteiro, 45
Mounjid, 122
Naumann, 173
Nika, 205
Nogueiras, 206
O’Sullivan, 221
Oliveira, 50
Olsen, 210
Oosterlee, 155, 161, 184, 200
Ortiz-Gracia, 195
Owens, 190
Pallavicini, 31
Peláez, 208
Petrov, 210
Pham, 124
Phillips, 178
Picarelli, 58
Pignotti, 51
Pironneau, 32
Pirvu, 67
Pitkin, 103
Polaski, 43
Possamai, 23
Pu, 211
Reisinger, 33
Reyners, 83
Rosenbaum, 122
Ruan, 224
Ruiz, 206
Sabbagh, 24
Saliba, 122
Salvador, 107
Santos, 45
Ševčovič, 107, 111, 112, 113
Sherris, 131
Shevchenko, 144
Silva, F., 58
Silva, P.M., 214
Smith, 60
Snoeijer, 216
Song, 24
Spanderen, 92
Statti, 55
Steffensen, 68
Stockinger, 57
Ströjby, 91
Suárez-Taboada, 51
Szücs, 114
Tempone, 143
Teng, 128
Thirurajah, 131
Touzi, 23, 34
Trnovská, 112
Trück, 223
van der Stoep, 161
Vanmaele, 56
Vassilovski, 217
Vázquez, 44, 51, 89, 107, 121
Venkataramanan, 221
Vilar, 208
Vives, 165, 201
von Wurstemberger, 86

Ware, 35
Wunderlich, 63

Yan, 223
Yu, 124

Zeron, 206
Zhang, W., 224
Zhang, Y., 33, 61
Zhu, 190
Ziveyi, 131, 169
Zubelli, 77