



HAL
open science

Optimal transport, geometry and Monte-Carlo methods for nonlinear PDEs: A ride in mathematical finance

Pierre Henry-Labordere

► **To cite this version:**

Pierre Henry-Labordere. Optimal transport, geometry and Monte-Carlo methods for nonlinear PDEs: A ride in mathematical finance. Probability [math.PR]. Université Paris Dauphine, 2014. tel-01088419

HAL Id: tel-01088419

<https://theses.hal.science/tel-01088419>

Submitted on 27 Nov 2014

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

CEREMADE - UMR 7534
UNIVERSITE PARIS-DAUPHINE

Mémoire d'habilitation à diriger des recherches en
Mathématiques appliquées

par

Pierre Henry-Labordère
Global Quantitative Research, Société Générale
CMAP, Ecole Polytechnique

Optimal Transport, Geometry and Monte-Carlo
methods for nonlinear PDEs:
A ride in mathematical finance

Soutenu le 3 Novembre 2014

Jury:

Coordinateur:

M. Bruno BOUCHARD, Professeur à l'Université Paris-Dauphine

Rapporteurs:

M. Peter FRIZ, Professeur à TU Berlin

M. Halil Mete SONER, Professeur à l'ETH Zurich

M. Denis TALAY, Professeur à l'INRIA Sophia Antipolis

Examineurs:

M. Yann BRENIER, Professeur à l'Ecole Polytechnique

Mme. Nicole EL KAROUI, Professeur à l'Université Pierre et Marie Curie

M. Nizar TOUZI, Professeur à l'Ecole Polytechnique

Contents

1	Martingale optimal transport	1
1.1	Introduction	1
1.2	Optimal transport in a nutshell	2
1.2.1	Fréchet-Hoeffding solution	4
1.2.2	Brenier's solution	5
1.2.3	Some symmetries	6
1.2.4	Multi-marginals and infinitely-many marginals case	7
1.2.5	Link with Hamilton-Jacobi equation	8
1.3	Martingale optimal transport [A2,A3,A4,S1,S2,S3,S4]	8
1.3.1	Dual formulation [A2]	10
1.3.2	A discrete martingale Fréchet-Hoeffding solution [S4]	11
1.3.3	Martingale Brenier's solution ♣	14
1.3.4	Symmetries in MOT [S4]	15
1.3.5	Martingale McCann's interpolation ♣	15
1.3.6	Multi-marginals extension [S4]	17
1.3.7	Infinitely-many marginals extension [S3]	18
1.4	Constrained martingale optimal transport	22
1.4.1	VIX constraints [S1]	23
1.4.2	Entropy penalty [A2]	25
1.4.3	American options ♣	26
1.5	Link with Skorohod embedding problem [A3, S2]	27
1.5.1	Probabilistic framework	27

1.5.2	Model-free super-hedging problem	27
1.5.3	Azema-Yor's solution revisited [A3]	29
1.5.4	Multi-marginals [S2]	31
1.6	Stochastic inequalities	32
1.6.1	Doob's inequality revisited	32
1.6.2	Burkholder-Davis-Gundy inequality revisited ♣	33
2	Asymptotics of implied volatility: A geometric approach	37
2.1	Motivation: risk model issues	37
2.2	From risk model issues to geometry [B1,A10,A11,A12]	38
2.2.1	Heat kernel on a Riemannian manifold	40
2.2.2	Riemann surfaces	42
2.2.3	Unified Asymptotic Implied Volatility [B1]	44
2.3	Large-strike behavior: Gaussian estimates ♣ [B1]	45
2.3.1	Time-homogeneous local volatility model	46
2.4	Solvable local and SVMs [A13]	49
2.4.1	Normal SABR	49
2.4.2	Superpotential and local volatility	50
3	Numerical solutions of non-linear PDEs	53
3.1	Some examples of non-linear PDEs	53
3.1.1	Uncertain volatility models	54
3.1.2	Local stochastic volatility models	55
3.1.3	Counterparty risk valuation	56
3.2	Second-order BSDEs and UVM [A8]	58
3.2.1	The algorithm	59
3.3	McKean SDEs and calibration in mathematical finance [A6,B2]	60
3.3.1	Malliavin's disintegration	61
3.3.2	The algorithm: toy example	61
3.4	Marked branching diffusions and CVA [A5,A1,B2]	62
3.4.1	Algorithm	66

<i>CONTENTS</i>	iii
3.5 Marked Malliavin branching diffusions ♣ [P4]	67
Publications	73

Introduction

Almost eleven years have passed through since my PhD in Theoretical Physics at Ecole Normale Supérieure in Paris under the supervision of Prof. Bernard Julia. My thesis was about dualities in M-theory and in supergravity theories involving various (recent) mathematical domains ranging from (real) algebraic geometry, super Kac-moody algebras to number automorphic forms. We have shown that the bosonic sector of supergravity theories, describing the low-energy limit of superstring theories/M-theory in ten and eleven dimensions, can be derived from a truncated Borchers super-algebra, extending the U-duality symmetries (see [PhD1, PhD2, PhD3]). These super-algebras pop up naturally as the Picard group of del Pezzo surfaces. These U-duality symmetries are a key tool to infer the strong coupling regime (inaccessible from perturbative string theories) of string theories (see [PhD4] for the computation of the effective Lagrangian of non-perturbative orientifold planes using U-symmetries, and automorphic forms).

After a short position at Imperial college in the Theoretical Physics department, I moved to quantitative finance, first at Barclays Capital in London and then to Société Générale in Paris. The present report is an overview of my research activity in an investment bank, mainly Société Générale. Although this research focuses and applies mainly to option pricing, I will emphasize the ramifications with numerous mathematical fields such as optimal transport, analysis on Riemannian geometry and numerical solutions of non-linear partial differential equations. New interesting mathematical questions emerge naturally from practical problems.

In comparison with the academic world, doing applied mathematics in industry requires producing (robust) numerical algorithms - the mathematical analysis coming at the end. As a conclusion, our research is strongly linked with numerical algorithms and all the subjects described below have lead to a numerical implementation. Some of them have been implemented at an industrial scale. In particular, the question of solving high-dimensional non-linear PDEs is a central theme, overviewed in Chapter 3.

We have decided to present our research activity in three (connected) parts (see my mathematical planisphere, Figure 1). Parts of these subjects are original, specially in the fields of mathematical finance. We can mention:

- Introduction of the martingale optimal transportation problem.
- A stochastic approach to the Skorohod embedding problem.
- New peacock processes.

- Risk-neutral weighted Monte-Carlo.
- Use of (geometrical) heat kernel expansion for deriving short-time asymptotics of implied volatility for (stochastic) volatility models.
- Use of 2-BSDE schemes for uncertain volatility models.
- Use of McKean particle method for solving various calibration issues in mathematical finance.
- Generalization of branching processes for solving some semi-linear PDEs and some non-Markovian BSDEs.
- Use of supersymmetries for the classification of solvable stochastic differential equations.

In Chapter 1, starting from the computation of model-independent bounds for exotic options, we introduce a new martingale version of the Monge-Kantorovich optimal transport problem (see [A2, A3, A4, S1, S2, S3, S4]). In this chapter, we will highlight the main differences with the optimal transport problem. This field has been popularized recently by various schools and conferences (including Polytechnique 2012, IHP 2013, Vienna 2013, Amsterdam 2014, ...). We outline also connections with the Skorohod embedding problem, for which we deduce new solutions in the case of multi-marginals, and path-wise (stochastic) inequalities. This topics will be covered in a forthcoming monograph “Martingale optimal transport” written with my collaborator Nizar Touzi from Ecole Polytechnique.

In Chapter 2, we summarize our work on asymptotic expansions of stochastic volatility models (in short SVM) (see [A7, A10, A11, A12, A13, B1]). The small-maturity behavior of the implied volatility is obtained using short-time heat kernel expansion on complete Riemannian manifolds. The large-strike behavior is studied by using Gaussian estimates of Schrödinger-like equations. This part is extensively developed in my first book [B1] “Analysis, Geometry, and Modeling in Finance: Advanced Methods in Option Pricing,” published by Chapman and Hall. Our toy model will be the SABR model which is connected to the geometry of the Poincaré plane. Various recent meetings (including Vienna 2009, Berlin 2011, London 2013 ...), with partial focus on small time asymptotic, perturbation theory, heat kernel methods, large deviations, and so on, have paid tribute to this development.

In Chapter 3, we focus on stochastic representations of second-order non-linear partial differential equations (see [B2, A1, A5, A6, A8, A9]). Due to the no-arbitrage condition in mathematical finance, we will treat exclusively the parabolic case. This subject is extensively covered in the book [B2] “Nonlinear Option Pricing” written with my former colleague Julien Guyon. We give a short overview of nonlinear PDEs arising in finance. As these PDEs suffer for the curse of dimensionality, their numerical solutions require probabilistic methods such as backward stochastic differential equations, particle method, and branching diffusions. The proposal numerical algorithms will be illustrated by various real-life problems such as pricing in uncertain volatility models, calibration of local stochastic volatility models, and counterparty valuation adjustment. Finally, we sketch a new method based on branching diffusions indexed by Malliavin’s weight, suitably *renormalized*, that lead to a *forward Monte-Carlo scheme* that can eventually tackled the resolution of fully nonlinear parabolic PDEs [P4].

We have tried to give a pedagogical presentation of our research topics motivated by our activity in investment banking. Technical details are reported in our research papers and in our books. In particular main results are often presented under some restrictive assumptions and illustrated by simple numerical examples. Readers can consult my research papers and books for detailed proofs and additional numerical examples. For further reference, published articles are cited with the index **A**, submitted papers with the index **S** and our books with the index **B**. Sections which present research proposals that we think deserve further analysis, eventually explored with PhD students, are marked with the symbol ♣.

Acknowledgements

First I would like to thank my main collaborator Nizar Touzi at Ecole Polytechnique for our fruitful collaborations on optimal transport and numerical solutions of nonlinear PDEs. I hope that we can continue to interact, eventually sharing PhD students together. I would like also to thank Bruno Bouchard, who kindly accepts to supervise my HDR habilitation, for his help and his support during the preparation of this thesis.

I want to express my deep gratitude to my *Rapporteurs* Peter Friz, Mete Soner and Denis Talay. It is a great honor to have them as referees. Special thanks also to Bruno Bouchard, Yann Brenier, Nicole El Karoui and Nizar Touzi.

I am grateful also to my colleagues at Société Générale for fruitful discussions, in particular L. Bergomi who gives me the opportunity to prepare this habilitation. I would like also to thank my former sparring partner J. Guyon for fruitful discussions, collaborations and friendship. Over the year, we have benefited from conversations with Marco Avellaneda, Matthias Beiglböck, Lorenzo Bergomi, Bruno Bouchard, Peter Carr, Bruno Dupire, Nicole El Karoui, Alfred Galichon, Emmanuel Gobet, Archil Gulisashvili, Benjamin Jourdain, Peter Laurence, Claude Martini, Friedrich Penkner, Denis Talay, Xiaolu Tan and Nizar Touzi.

Last but not least, I would like to thank my family Véronique, Aurélien, Emma, Florian, who endure patiently in the same year the writing of my book [B2] and this habilitation.

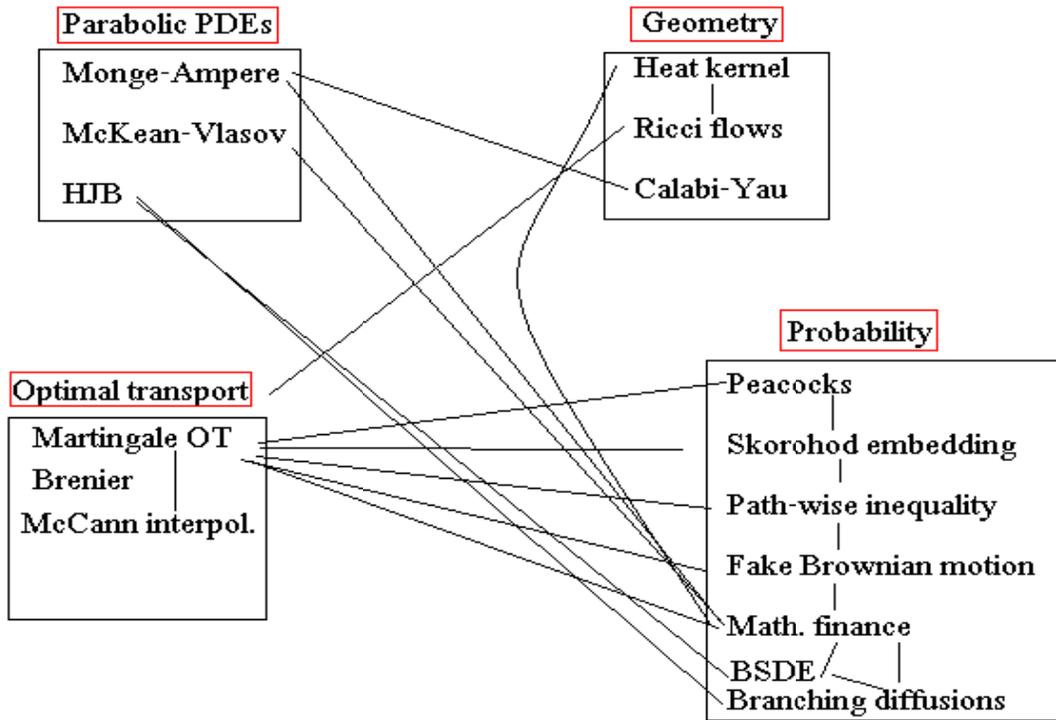


Figure 1: My HDR mathematical planisphere.

Chapter 1

Martingale optimal transport

In this chapter, we present our research on a new constrained version of the optimal transport. We will try to highlight the new differences between the optimal transport and its martingale counterpart and motivate this topic in the context of mathematical finance. The martingale optimal transport that we introduced in an internal publication in 2010 was substantially developed with our coworkers at Ecole Polytechnique and Vienna university. Our contribution is two-fold: On a practical side, it allows to better understand payoffs that can be well-hedged by Vanilla options. On a theoretical side, martingale optimal transport leads to new results in probability and mathematical finance: new solutions to the multi-marginals Skorohod embedding problem, a new approach for deriving path-wise inequalities, new peacock processes and robust hedging. This subject has attracted much attention: various summer schools and workshops (Polytechnique <http://www.cmap.polytechnique.fr/~euroschoollmathfi12/>, ETH, Vienna, Amsterdam) have been devoted to it. Also various reading groups in Mathematical Finance (THU Berlin, Imperial College <http://www3.imperial.ac.uk/mathfin/events/readinggroup>, ...) on the applications of this new optimal transport have been created.

This chapter originates for our research papers [A2, A3, A4, S1, S2, S3, S4]. In what follows, I assume for the sake of simplicity a zero interest rate. Although optimal transport is now developed in various monograph (see e.g [107]), we present in the next section a short introduction highlighting the main features that will be generalized in the martingale version.

1.1 Introduction

Optimal transport, first introduced by G. Monge in his work “Théorie des déblais et des remblais” (1781), has recently generated widely interest in various mathematical domains as highlighted by the last Fields medallist C. Villani (2010). Let us mention the analysis of non-linear (kinetic) partial differential equations arising in statistical physics such as McKean-Vlasov PDE, infinite-dimensional linear programming, mean-field limits, convergence of particle methods and study of Ricci flows in differential geometry.

Despite its numerous ramifications in analysis and probability, optimal transport had not yet attracted the attention of academics/practitioners in financial mathematics. However, various long-standing problems in quantitative finance can be tackled using the framework of optimal transport: calibration of (hybrid) models on market smiles, arbitrage-free construction of smiles, computation of efficient model-independent bounds for exotic options, ...

The main problem, when evaluating an exotic option, is how to choose an “appropriate” pricing model, a model being characterized by a martingale measure. “Appropriate” means here that the model allows to capture the main risks of the exotic option under consideration: At-the-money volatility, skew, forward volatility, forward skew,... One may impose that the model is calibrated to a set of (liquid) market instruments or matches some historical levels. Because Vanilla options are liquid, hence the most suitable hedge instruments, the model has to comply with their market prices. In mathematical terms, the marginals of the probability measure, for a discrete set of dates are given. Only a few models such as the Dupire local volatility [41] or (multi-factor) local stochastic volatility models [A7, A9] can achieve efficient calibration to Vanilla options.

We follow a different route. Instead of postulating a model, we focus on the computation of model-independent bounds consistent with Vanillas, eventually additional instruments such as VIX futures. By duality, we will show that these bounds are attained by some arbitrage-free models. The computation of model-independent bounds for exotic options can then be framed as a constrained optimal transportation problem. The dual Kantorovich-like formulation can be interpreted as a robust sub/super-hedging strategy.

This problem will be also linked with the Skorohod embedding problem which consists in building a stopping time τ such that a Brownian motion stopped at that time has the same law than a probability \mathbb{P}^T : $B_\tau \sim \mathbb{P}^T$. In option pricing layman’s term, by the Dambis-Dubins-Schwarz time change theorem, it consists in building an arbitrage-free model consistent with T -Vanillas for some maturity date T . Finally, we show that the martingale optimal framework is an efficient tool for deriving path-wise (stochastic) inequalities.

1.2 Optimal transport in a nutshell

As our approach in martingale optimal transport follows closely the analysis in classical optimal transport, we review in this section key notions in optimal transport such as Monge-Kantorovich duality and Brenier’s theorem. We give its interpretation in mathematical finance. As a consequence, instead of presenting the main results in the setup of Polish spaces X , we focus on the simple case $X = \mathbb{R}_+$, eventually \mathbb{R}^d . Our main reference is Chapter 1, 2, 3 and 5 in [107].

Let us start with two assets S_1 and S_2 evaluated at the same maturity date T and consider a payoff $c(s_1, s_2)$ depending on these two underlyings. c is an upper semi-continuous function with linear growth. We assume that T -Vanilla options on each asset are traded on the market for all strikes. We can then deduce the T -marginal distributions

$$\mathbb{P}^i(K) = \partial_K^2 C^i(K), \quad i = 1, 2$$

where $C^i(K)$ is the market value of a call option with strike K and maturity T on asset i . The second-order derivative is defined almost everywhere as $C^i(K)$ is a convex function. Below we

denote S_1 (resp. S_2) the random variable such that $S_1 \sim \mathbb{P}^1$ (resp. $S_2 \sim \mathbb{P}^2$). We assume that the first moment of \mathbb{P}^1 (resp. \mathbb{P}^2) is finite equal to the spot value at $t = 0$: S_1^0 (resp. S_2^0).

The model-independent super-replication price is then defined as

$$\text{MK}_2 = \inf_{\mathcal{P}^*(\mathbb{P}^1, \mathbb{P}^2)} \mathbb{E}^{\mathbb{P}^1}[u_1(S_1)] + \mathbb{E}^{\mathbb{P}^2}[u_2(S_2)] \quad (1.1)$$

where $\mathcal{P}^*(\mathbb{P}^1, \mathbb{P}^2)$ is the set of all measurable functions $(u_1, u_2) \in L^1(\mathbb{P}^1) \times L^1(\mathbb{P}^2)$ such that

$$u_1(s_1) + u_2(s_2) \geq c(s_1, s_2) \quad (1.2)$$

for \mathbb{P}^1 -almost all $s_1 \in \mathbb{R}_+$ and \mathbb{P}^2 -almost all $s_2 \in \mathbb{R}_+$. This static super-replication strategy consists in holding European payoffs $u_1(s_1)$ and $u_2(s_2)$ with $(t = 0)$ market prices $\mathbb{E}^{\mathbb{P}^1}[u_1(S_1)]$, $\mathbb{E}^{\mathbb{P}^2}[u_2(S_2)]$ such that the intrinsic value of this portfolio $u_1(s_1) + u_2(s_2)$ is greater than the payoff $c(s_1, s_2)$. In practice, these European payoffs can be replicated by holding a strip of put/call T -Vanillas through the Taylor expansion:

$$u_1(s_1) = u_1(s_1^0) + u_1'(s_1^0)(s_1 - s_1^0) + \int_0^{s_1^0} u_1''(K)(K - s_1)^+ dK + \int_{s_1^0}^{\infty} u_1''(K)(s_1 - K)^+ dK$$

Derivatives are understood in the distribution sense.

Optimal transport formulation

The linear program (1.1), called the Monge-Kantorovich formulation of optimal transport, can be dualized by introducing a (positive) Kuhn-Tucker multiplier (i.e., a probability measure on \mathbb{R}_+^2) associated to the inequality (1.2):

$$\text{MK}_2 = \sup_{\mathbb{P} \in \mathcal{P}(\mathbb{P}^1, \mathbb{P}^2)} \mathbb{E}^{\mathbb{P}}[c(S_1, S_2)] \quad (1.3)$$

with $\mathcal{P}(\mathbb{P}^1, \mathbb{P}^2) = \{\mathbb{P} : S_1 \overset{\mathbb{P}}{\sim} \mathbb{P}^1, S_2 \overset{\mathbb{P}}{\sim} \mathbb{P}^2\}$. This dual expression coincides with the Kantorovich formulation of optimal transport. This result is known as the Kantorovich duality. Note that the optimal transport is usually presented in textbooks with an inf instead of a sup.

The optimization MK_2 consists therefore in maximizing the cost function $\mathbb{E}^{\mathbb{P}}[c(S_1, S_2)]$ over the (convex) set of joint measures \mathbb{P} with marginals \mathbb{P}^1 and \mathbb{P}^2 . For an upper semi-continuous cost function, as the set $\mathcal{P}(\mathbb{P}^1, \mathbb{P}^2)$ is weakly compact from Prokhorov's theorem, the maximum is attained. Moreover, the value of the infimum does not change in the right-hand side of (1.1) if one restricts the definition of $\mathcal{P}^*(\mathbb{P}^1, \mathbb{P}^2)$ to functions (u_1, u_2) which are bounded and continuous. For further reference, we quote the following theorem that proves that the infimum is attained by a pair (u, u^*) of lower semi-continuous c -convex functions

$$\text{MK}_2 = \inf_{u \in C_b} \mathbb{E}^{\mathbb{P}^1}[u^*(S_1)] + \mathbb{E}^{\mathbb{P}^2}[u(S_2)] \quad (1.4)$$

with $u^*(s_1) = \sup_{s_2 \in \mathbb{R}_+} \{c(s_1, s_2) - u(s_2)\}$.

Note that all references in optimal transport present first the primal (1.3) and then the dual formulation (1.1). In the context of mathematical finance, the dual (1.1), being interpreted as the robust upper bound of the fair value of c , has a clear financial interpretation and therefore is presented first.

Below, we give an example of applications of the (classical) optimal transport in the construction of arbitrage-free currency cross-rate smiles, drawn from [A2].

Example 1.2.1 (Illiquid currency cross-rate smile) *We consider the arbitrage-freeness of a currency cross-rate smile using the smiles of two related exchange rates. More precisely, let us consider Vanilla payoffs on the exchange rates linking the following three currencies: JPY, US dollars and Australian dollars. We denote the dollar price of one Japanese Yen at maturity T by $S_1 \equiv S_T^{\$/\text{JPY}}$ and likewise the dollar price of one Australian dollar by $S_2 \equiv S_T^{\$/\text{AUD}}$. The value of one JPY in Australian dollars is then given by $S_3 \equiv S_T^{\text{AUD}/\text{JPY}} = S_2/S_1$. Hence, a cross-rate call option is equivalent to an option to exchange one asset for the other asset and its payoff equals $(S_2 - KS_1)^+$. The triangle consisting in T -Vanillas on S_1 , S_2 and S_3 is arbitrage-free if the bound MK_3 (≤ 0) defined below is zero:*

$$\text{MK}_3 = \inf_{\mathcal{P}^*(\mathbb{P}^1, \mathbb{P}^2, \mathbb{P}^3)} \mathbb{E}^{\mathbb{P}^1}[u_1(S_1)] + \mathbb{E}^{\mathbb{P}^2}[u_2(S_2)] + \mathbb{E}^{\mathbb{P}^3}[S_1 u_3(S_3)]$$

where $\mathcal{P}^*(\mathbb{P}^1, \mathbb{P}^2, \mathbb{P}^3)$ is the set of continuous functions on $L^1(\mathbb{P}^1) \times L^1(\mathbb{P}^2) \times L^1(\mathbb{P}^3)$ such that

$$u_1(s_1) + u_2(s_2) + s_1 u_3(s_2/s_1) \geq 0$$

for \mathbb{P}^1 -almost all $s_1 \in \mathbb{R}_+$, \mathbb{P}^2 -almost all $s_2 \in \mathbb{R}_+$. Indeed if $\text{MK}_3 < 0$, then the arbitrage can be locked in by buying the static portfolio $u_1(s_1) + u_2(s_2) + s_1 u_3(s_2/s_1)$ at the price MK_3 . MK_3 is equivalent to (see Equation 1.4)

$$\text{MK}_3 = \inf_{u_1 \in L^1(\mathbb{P}^1), u_2 \in L^1(\mathbb{P}^2)} \mathbb{E}^{\mathbb{P}^1}[u_1(S_1)] + \mathbb{E}^{\mathbb{P}^2}[u_2(S_2)] - \mathbb{E}^{\mathbb{P}^3}[S_1 u_3^*(S_3)]$$

with u_3^* is the (multiplicative) inf-convolution of u_1 and u_2 :

$$u_3^*(s_3) \equiv \inf_{s_1 \in \mathbb{R}_+} \frac{1}{s_1} \{u_1(s_1) + u_2(s_3 s_1)\}$$

MK_3 defines a linear program that can be solved with a simplex algorithm. A triangle arbitrage is detected if $\text{MK}_3 \neq 0$, meaning by duality that the subset of probability measures on \mathbb{R}_+^2 , $\mathcal{P}(\mathbb{P}^1, \mathbb{P}^2, \mathbb{P}^3) \equiv \{\mathbb{P} : S_1 \stackrel{\mathbb{P}}{\sim} \mathbb{P}^1, S_2 \stackrel{\mathbb{P}}{\sim} \mathbb{P}^2, \int x^2 \mathbb{P}(x, Kx) dx = \mathbb{P}^3(K) \forall K \in \mathbb{R}_+\}$, is empty.

1.2.1 Fréchet-Hoeffding solution

Under the so-called Spence-Mirrlees condition, $c_{12} \equiv \partial_{s_1 s_2} c > 0$, the optimal transport (1.3) can be solved explicitly. Let F_1, F_2 denote the cumulative distribution functions of \mathbb{P}^1 and \mathbb{P}^2 . For the sake of simplicity, we will assume that \mathbb{P}^1 does not give mass to points and $c \in C^2$. The optimal measure \mathbb{P}^* has the form

$$\mathbb{P}^*(ds_1, ds_2) = \delta_{T(s_1)}(ds_2) \mathbb{P}^1(ds_1)$$

with T the push-forward of the measure \mathbb{P}^1 onto \mathbb{P}^2 : $T(x) = F_2^{-1} \circ F_1(x)$. The optimal upper bound is given by

$$\text{MK}_2 = \int_0^1 c(F_1^{-1}(u), F_2^{-1}(u)) du$$

This optimal bound can be attained by a static hedging strategy consisting in holding European payoffs $u_1(s_1) \in L^1(\mathbb{P}^1)$, $u_2(s_2) \in L^1(\mathbb{P}^2)$ with market prices $\mathbb{E}^{\mathbb{P}^1}[u_1(S_1)]$ and $\mathbb{E}^{\mathbb{P}^2}[u_2(S_2)]$:

$$\text{MK}_2 = \mathbb{E}^{\mathbb{P}^1}[u_1(S_1)] + \mathbb{E}^{\mathbb{P}^2}[u_2(S_2)]$$

with

$$u_2(x) = \int_0^x c_2(T^{-1}(y), y) dy, \quad u_1(x) = c(x, T(x)) - u_2(T(x))$$

The intrinsic value of this static European portfolio super-replicates the payoff at maturity:

$$u_1(s_1) + u_2(s_2) \geq c(s_1, s_2), \quad \forall (s_1, s_2) \in \mathbb{R}_+ \times \mathbb{R}_+$$

T is referred as the Brenier map (or Fréchet-Hoeffding) pushing forward \mathbb{P}^1 to \mathbb{P}^2 . This solution coincides with the Monge solution where the Monge optimal transport is by definition given by

$$M_2 = \sup_{\mathcal{P}^{\text{Monge}}(\mathbb{P}^1, \mathbb{P}^2)} \mathbb{E}^{\mathbb{P}}[c(S_1, S_2)]$$

with the set $\mathcal{P}^{\text{Monge}}(\mathbb{P}^1, \mathbb{P}^2) = \{\mathbb{P}_T \in \mathcal{P}(\mathbb{P}^1, \mathbb{P}^2) : \mathbb{P}_T(ds_1, ds_2) = \delta_{T(s_1)}(ds_2) \mathbb{P}^1(s_1)\}$. Note that this problem is more involved as the constraint on T is nonlinear (see section on Monge-Ampère equation). It is therefore simpler to start with the (relaxed) Monge-Kantorovich formulation and then show that it coincides with the Monge problem.

Example 1.2.2 (Basket option, $c(s_1, s_2) = (s_2 + s_1 - K)^+$) By applying Fréchet-Hoeffding solution, the upper bound for a two-asset basket option is attained by

$$\begin{aligned} u_2(x) &= (x - \bar{y})^+ \\ u_1(x) &= (T(x) + x - K)^+ - (T(x) - \bar{y})^+ \\ T(x) &= F_2^{-1} \circ F_1(x) \end{aligned}$$

with \bar{y} defined by $T(K - \bar{y}) = \bar{y}$.

1.2.2 Brenier's solution

The Fréchet-Hoeffding solution has been generalized in the case \mathbb{R}^d by Brenier [22] first in the case of a quadratic cost function and then extended to concave payoff $c = c(s_1 - s_2)$ by Gangbo and McCann [55] and others:

Theorem 1.2.3 (Brenier - $c(s_1, s_2) = -|s_1 - s_2|^2/2$) *If \mathbb{P}^1 has no atoms, then there is a unique optimal \mathbb{P}^* , which is a Monge solution:*

$$\mathbb{P}^* = \delta_{T(s_1)}(ds_2)\mathbb{P}^1(s_1)$$

with $T = \nabla u$. ∇u is the unique gradient of a convex function which pushes forward \mathbb{P}^1 to \mathbb{P}^2 : $\nabla u_{\#}\mathbb{P}^1 = \mathbb{P}^2$.

Note that this theorem has been generalized to a strictly concave, superlinear cost function $c(s_1, s_2) = c(s_1 - s_2)$. The Brenier map is then

$$T(x) = x - \nabla c^*(\nabla u)$$

for some c -convex function u which is uniquely fixed by the requirement $T_{\#}\mathbb{P}^1 = \mathbb{P}^2$. c^* is the Legendre transform of c . If we assume that \mathbb{P}^1 and \mathbb{P}^2 are absolutely continuous with respect to the Lebesgue measure, $\mathbb{P}^1(dx) \equiv p_1(x)dx$, $\mathbb{P}^2(dx) \equiv p_2(x)dx$, and $T \in C^1$, this implies that u is then solution of a Monge-Ampère equation:

$$\det(1 - \nabla(\nabla c)^*(\nabla u)) = \frac{p_1(x)}{p_2(T(x))}$$

Note that this nonlinear PDE arises in mathematical finance - see for example the problem of quantile hedging [25] - through the following stochastic representation

$$\sup_{\sigma \in S_d^+, \det(\sigma)=1} \text{Tr}[\sigma \nabla^2 u] + f(x) = 0$$

where S_d^+ denotes the set of nonnegative symmetric d -dimensional matrices. This corresponds to an (infinite-horizon) HJB control problem in an multi-asset uncertain volatility models (see Chapter 3).

1.2.3 Some symmetries

Spence-Mirrlees condition

The Spence-Mirrlees condition is very natural from a financial point of view. If we shift the payoff c by some European payoffs $U_1 \in L^1(\mathbb{P}^1)$, $U_2 \in L^1(\mathbb{P}^2)$:

$$\bar{c}(s_1, s_2) = c(s_1, s_2) + U_1(s_1) + U_2(s_2)$$

then the Monge-Kantorovich bound for \bar{c} should be

$$\text{MK}_2(\bar{c}) = \text{MK}_2(c) + \mathbb{E}^{\mathbb{P}^1}[U_1(S_1)] + \mathbb{E}^{\mathbb{P}^2}[U_2(S_2)]$$

The payoff \bar{c} is precisely invariant under the Spence-Mirrlees condition : $\bar{c}_{12} = c_{12}$.

Mirror coupling: co-monotone rearrangement map

Similarly, the upper bound under the condition $c_{12} < 0$ is attained by the co-monotone rearrangement map

$$T(s_1) = F_2^{-1} \circ (1 - F_1(-s_1))$$

This can be obtained by applying the parity transformation $\mathcal{P}(s_1, s_2) = (-s_1, s_2)$. For each measure \mathbb{P} matching the marginals \mathbb{P}^1 and \mathbb{P}^2 , we can associate the measure $\mathcal{P}_*\mathbb{P}$ matching the marginals $\mathcal{P}_*\mathbb{P}^1$ and \mathbb{P}^2 with cumulative distributions $1 - F_1(-s_1)$ and $F_2(s_2)$. We conclude as the Monge-Kantorovich bounds for c and $\tilde{c}(s_1, s_2) \equiv c(-s_1, s_2)$ coincides as $\mathbb{E}^{\mathbb{P}}[c] = \mathbb{E}^{\mathcal{P}_*\mathbb{P}}[\tilde{c}]$. Similarly, by replacing c by $-c$, we obtain that the co-monotone rearrangement map gives the *lower* bound under the condition $c_{12} > 0$.

1.2.4 Multi-marginals and infinitely-many marginals case

Most of the literature on optimal transport focuses on the 2-asset case. For applications in mathematical finance, it could be interesting to study the case of a Vanilla multi-asset payoff $c(s_1, \dots, s_n)$ depending on n assets evaluated at the same maturity date. We define the n -asset optimal transportation problem (by duality) as

$$\text{MK}_n = \sup_{\mathbb{P} \in \mathcal{P}(\mathbb{P}^1, \dots, \mathbb{P}^n)} \mathbb{E}^{\mathbb{P}}[c(S_1, \dots, S_n)]$$

with $\mathcal{P}(\mathbb{P}^1, \dots, \mathbb{P}^n) = \{\mathbb{P} : S_i \stackrel{\mathbb{P}}{\sim} \mathbb{P}^i, \forall i = 1, \dots, n\}$. This problem has been studied by Gangbo [54] and recently by Carlier [29] (see also Pass [98]) with the following requirement on the payoff:

Definition 1.2.4 (see [29]) $c \in C^2$ is strictly monotone of order 2 on \mathbb{R}_+^n if for all $(i, j) \in \{1, \dots, n\}^2$ with $i \neq j$, all second order derivatives $\partial_{ij}c$ are positive.

We have

Theorem 1.2.5 (see [54, 29]) *If c is strictly monotone of order 2, there exists a unique optimal transference plan for the MK_n transportation problem, and it has the form*

$$\mathbb{P}^*(ds_1, \dots, ds_2) = \mathbb{P}^1(ds_1) \prod_{i=1}^n \delta_{T^i(s_1)}(ds_i), \quad T^i(s_1) = F_i^{-1} \circ F_1(s_1), \quad i = 2, \dots, n$$

The optimal upper bound is

$$\text{MK}_n = \int c(x, T_2(x), \dots, T_n(x)) \mathbb{P}^1(dx)$$

An extension to the infinite many marginals case has been obtained recently by Pass [99] who studies

$$\text{MK}_\infty = \sup_{\mathbb{P} : X_t \stackrel{\mathbb{P}}{\sim} \mathbb{P}^t, \forall t \in (0, 1]} \mathbb{E}\left[h\left(\int_0^1 X_t dt\right)\right]$$

where h is a convex function. Let F_t the cumulative distribution of \mathbb{P}^t . Define the stochastic process $X_t^{\text{opt}}(\omega) = F_t^{-1}(\omega)$, $\omega \in [0, 1]$. The underlying probability space is the interval $[0, 1]$ with Lebesgue measure. We have

Theorem 1.2.6 (see [99]) *The process X_t^{opt} is the unique maximizer in MK_∞ .*

1.2.5 Link with Hamilton-Jacobi equation

Here we take a cost function $c(s_1, s_2) = L(s_1 - s_2)$ with L a strictly concave function such that the Spence-Mirrlees condition is satisfied. From formulation (1.4), one can link the Monge-Kantorovich formulation to the solution of an Hamilton-Jacobi equation through the Hop-Lax formula:

Proposition 1.2.7 (see e.g. [107])

$$\text{MK}_2 = \inf -\mathbb{E}^{\mathbb{P}^1}[u(0, S_1)] + \mathbb{E}^{\mathbb{P}^2}[u(1, S_2)]$$

where the supremum is taken over all continuous viscosity solutions u to the following HJ equation:

$$\partial_t u(t, x) + H(Du) = 0, \quad H(p) \equiv \inf_q \{pq - L(q)\} \quad (1.5)$$

In the next section, we introduce a martingale version of optimal transport, first developed in [A3, A4] where we have obtained a Monge-Kantorovich duality result. In [S2, S3, S4], we explicitly solve the martingale optimal transport for a large class of cost functions. A constrained martingale version is presented in [S1] and in [A2] we focus on numerical implementations. We briefly summarized the main results below.

1.3 Martingale optimal transport [A2,A3,A4,S1,S2,S3,S4]

Here, we consider a payoff $c(s_1, s_2)$ depending on a *single* asset evaluated at two dates $t_1 < t_2$. As above, we assume that Vanilla options of all strikes with maturities t_1 and t_2 are traded and therefore, one can imply the distribution of S_1 and S_2 (here $S_1 \equiv S_{t_1}, S_2 \equiv S_{t_2}$). The model-independent upper bound, consistent with t_1 and t_2 Vanilla options, can then be defined as a martingale version of the above optimal transport problem MK_2 :

$$\widetilde{\text{MK}}_2 \equiv \inf_{\mathcal{M}^*(\mathbb{P}^1, \mathbb{P}^2)} \mathbb{E}^{\mathbb{P}^1}[u_1(S_1)] + \mathbb{E}^{\mathbb{P}^2}[u_2(S_2)]$$

where $\mathcal{M}^*(\mathbb{P}^1, \mathbb{P}^2)$ is the set of measurable functions $u_1 \in L^1(\mathbb{P}^1), u_2 \in L^1(\mathbb{P}^2)$ and h a bounded continuous function on \mathbb{R}_+ such that

$$u_1(s_1) + u_2(s_2) + h(s_1)(s_2 - s_1) \geq c(s_1, s_2) \quad (1.6)$$

for \mathbb{P}^1 -almost all $s_1 \in \mathbb{R}_+$ and \mathbb{P}^2 -almost all $s_2 \in \mathbb{R}_+$. This corresponds to a semi-static hedging strategy which consists in holding European payoffs u_1 and u_2 and applying a delta strategy at t_1 ,

generating a P&L $h(s_1)(s_2 - s_1)$ at t_2 with zero cost. Note that in comparison with the optimal transport $\widetilde{\text{MK}}_2 \leq \text{MK}_2$ due to the appearance of the function h .

At this point, a natural question is how the classical results in optimal transport generalize in the present martingale version. In particular, the quadratic cost which is the main example in optimal transport is here degenerate (see below). We follow closely our introduction of optimal transport and explain how the various concepts previously explained extend to the present setting. Our research partly originates from a systematic derivation of Skorohod embedding solutions and particle methods for non-linear McKean stochastic differential equations appearing in the calibration of financial models (see Chapter 3). From a practical point of view, the derivation of these optimal bounds allows to better understand the risk of exotic options as illustrated in the next example.

Example 1.3.1 (Forward-start options, see A2) *We consider model-independent bounds for forward-start options with payoff $(s_2/s_1 - K)^+$. We assume that we are calibrated to Eurostock implied volatilities (pricing date = 2 Feb. 2010) at time $t_1 = 1$ year and $t_2 = 1.5$ years with $N = 18$ strikes $\in [30\%, 200\%]$. The dual reads as (with $K_1^0 = K_2^0 = 0$ corresponding to forward prices)*

$$\min_{\nu, (\omega_1^j), (\omega_2^j), h(\cdot)} \nu + \sum_{j=0}^N \omega_1^j C(t_1, K_1^j) + \sum_{j=0}^N \omega_2^j C(t_2, K_2^j)$$

$$F(s_1, s_2) \geq \left(\frac{s_2}{s_1} - K \right)^+, \quad \forall (s_1, s_2) \in (\mathbb{R}_+)^2$$

with $F \equiv \sum_{j=0}^N \omega_1^j (s_1 - K_j)^+ + \sum_{j=0}^N \omega_2^j (s_2 - K_j)^+ + \nu + h(s_1)(s_2 - s_1)$ and $C(t_i, K)$ the market value of a call of maturity t_i and K . We have compared the upper and lower bounds against prices produced by various (stochastic) volatility models commonly used by practitioners (see Fig. 1.1): Local volatility model [41], Bergomi model [19] which is a two-factor variance curve model and the local Bergomi model [A6,A9] which has the property that it is calibrated to the Vanilla smile. The LV and local Bergomi models have been calibrated to the Eurostock implied volatility (pricing date = 2 Feb. 2010). The Bergomi model has been calibrated to the variance-swap term structure.

We should emphasize that lower and upper bounds for each strike K correspond to a different martingale measure (see below our duality result). This is not the case if we do not include the martingality constraint as the upper/lower bounds are attained by the Fréchet-Hoeffding bounds for each strike.

For $K = 1$, the difference between the implied volatility given by the upper bound and the Bergomi model is around 10%. This result shows that forward-start options are poorly constrained by Vanilla smiles. As a conclusion, the (common) practise to calibrate stochastic volatility model on Vanilla smiles to price exotic options (depending strongly on forward/skew volatility) is inappropriate: Forward-start options are poorly hedged by t_1 and t_2 Vanilla options.

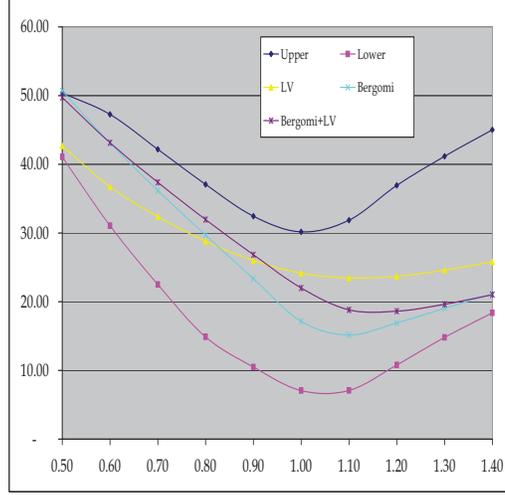


Figure 1.1: Implied volatility for cliquet options. Lower/Upper bounds versus (local) Bergomi and LV models for cliquet options (quoted in Black-Scholes volatility $\times 100$). Parameters for the Bergomi model: $\sigma = 2.0$, $\theta = 22.65\%$, $k_1 = 4$, $k_2 = 0.125$, $\rho = 34.55\%$, $\rho_{SX} = -76.84\%$, $\rho_{SY} = -86.40\%$.

1.3.1 Dual formulation [A2]

In [A4], we have established a dual version. We set

$$\widetilde{\text{MK}}_2^* = \sup_{\mathbb{P} \in \mathcal{M}(\mathbb{P}^1, \mathbb{P}^2)} \mathbb{E}^{\mathbb{P}}[c(S_1, S_2)] \quad (1.7)$$

where $\mathcal{M}(\mathbb{P}^1, \mathbb{P}^2) = \{\mathbb{P} : \mathbb{E}^{\mathbb{P}}[S_2|S_1] = S_1, S_1 \stackrel{\mathbb{P}}{\sim} \mathbb{P}^1, S_2 \stackrel{\mathbb{P}}{\sim} \mathbb{P}^2\}$ denotes the set of (discrete) martingale measures on \mathbb{R}_+^2 with marginals \mathbb{P}^1 and \mathbb{P}^2 . As $\mathcal{M}(\mathbb{P}^1, \mathbb{P}^2)$ is convex and weakly compact (see [A4]), the dual is attained by an extremal point. In principle from Krein-Milman's theorem, the space $\mathcal{M}(\mathbb{P}^1, \mathbb{P}^2)$ can be reconstructed from its extremal points (that can be obtained by looking at a specific cost function).

The set $\mathcal{M}(\mathbb{P}^1, \mathbb{P}^2)$ is feasible (i.e., non-empty) if and only if $\mathbb{P}^1, \mathbb{P}^2$ have the same mean S_0 and $\mathbb{P}^1 \leq \mathbb{P}^2$ are in convex order (see Kellerer, [83], and section 2.2 in [12]) meaning that

$$\mathbb{E}^{\mathbb{P}^1}[(S_1 - K)^+] \leq \mathbb{E}^{\mathbb{P}^2}[(S_2 - K)^+], \quad \forall K \in \mathbb{R}_+$$

In financial term, we exclude calendar spread arbitrage opportunities (on call options). The necessary condition can easily be obtained by applying the Jensen inequality on $(S_2 - K)^+$. The sufficient condition can be deduced by explicitly building a martingale measure matching marginals \mathbb{P}^1 and \mathbb{P}^2 from the Dupire local volatility model (see e.g. [70]).

Assum(c): Let $c : \mathbb{R}_+^2 \rightarrow [-\infty, \infty)$ be an upper semi-continuous function such that

$$c^+(s_1, s_2) \leq K \cdot (1 + s_1 + s_2) \quad (1.8)$$

on $(\mathbb{R}_+)^2$ for some constant K .

Theorem 1.3.2 (see A4) *Assume that $\mathbb{P}^1 \leq \mathbb{P}^2$ are Borel probability measures on \mathbb{R}_+ with first moments S_0 and assumption **Assum(c)** holds. Then there is no duality gap, i.e. $\widetilde{\text{MK}}_2^* = \widetilde{\text{MK}}_2$. Moreover, the primal value $\widetilde{\text{MK}}_2^*$ is attained, i.e. there exists a martingale measure $\mathbb{P}^* \in \mathcal{M}(\mathbb{P}^1, \mathbb{P}^2)$ such that $\widetilde{\text{MK}}_2^* = \mathbb{E}^{\mathbb{P}^*}[c]$. In general, the dual is not attained (see [A4] for a counterexample).*

The function h is interpreted as the Lagrange multiplier associated to the martingale condition. This result has generated new research in the field of robust hedging (see e.g. [2, 5, 12, 13, 14, 26, 38, 39, 106]). In [A3], we have generalized our duality result in a continuous-time setup. We also mention [38].

As previously mentioned, note that the quadratic cost, $c(s_1, s_2) = (s_2 - s_1)^2$ is degenerate in the present martingale version as

$$\mathbb{E}^{\mathbb{P}}[(S_2 - S_1)^2] = \mathbb{E}^{\mathbb{P}^2}[S_2^2] - \mathbb{E}^{\mathbb{P}^1}[S_1^2], \quad \mathbb{P} \in \mathcal{M}(\mathbb{P}^1, \mathbb{P}^2)$$

meaning that the cost value is identical for all $\mathbb{P} \in \mathcal{M}(\mathbb{P}^1, \mathbb{P}^2)$. In OT, one can prove that the minimization can be restricted to the class of c -convex function (see equation 1.4). A similar result holds in MOT where the c -convex property is replaced by the concave envelope:

Proposition 1.3.3 (see A4) *Assume that $\mathbb{P}^1 \leq \mathbb{P}^2$ are Borel probability measures on \mathbb{R}_+ with first moments S_0 and assumption **Assum(c)** holds. Then*

$$\widetilde{\text{MK}}_2^* = \sup_{u \in L^1(\mathbb{P}^2)} \mathbb{E}^{\mathbb{P}^1}[(c(S_1, \cdot) - u(\cdot))^{**}(S_1)] + \mathbb{E}^{\mathbb{P}^2}[u(S_2)] \quad (1.9)$$

where for a function g , g^{**} denotes its concave envelope, i.e., the smallest concave function greater or equal to g .

1.3.2 A discrete martingale Fréchet-Hoeffding solution [S4]

In this section (see [S4] for details), we solve explicitly $\widetilde{\text{MK}}_2$ under the martingale Spence-Mirrlees condition $\partial_{s_1 s_2 s_2} c > 0$. This gives a martingale measure, similar in spirit to the Fréchet-Hoeffding solution: Under the new condition $\partial_{s_1 s_2 s_2} c > 0$, the optimal measure is payoff-independent and depends only on the marginals \mathbb{P}^1 and \mathbb{P}^2 . Moreover, we will show that this condition guarantees that the dual infimum is attained which was not guaranteed from our previous duality result. The optimal measure is no more supported on a single map T as it was for Fréchet-Hoeffding's solution. Indeed, the martingale constraint can not be fulfilled in this case (except in the trivial case $\mathbb{P}^1 = \mathbb{P}^2$):

$$\mathbb{E}^{\mathbb{P}^*}[(S_2 - S_1) | S_1 = s_1] = T(s_1) - s_1 \neq 0$$

for

$$\mathbb{P}^*(ds_1, ds_2) = \delta_{T(s_1)}(ds_2)\mathbb{P}^1(ds_1), \quad T(x) = F_2^{-1} \circ F_1(x)$$

As a straightforward guess, we will assume that \mathbb{P}^* is supported along two maps $s_2 = T_d(s_1)$ and $s_2 = T_u(s_1)$. This intuition comes from [72] where the authors consider the problem of finding the optimal upper bound on the price of a forward-start straddle $|s_2 - s_1|$ and from an old result of Dubins, Schwarz [40] that characterizes the extreme points of the space of all distributions of discrete martingales as those that possess these two properties : (a) S_0 is fixed and (b) the conditional distribution of each S_k given the past up to time $k - 1$ is almost surely a two-valued distribution. This result should be put in light with the fact that a pricing binomial model is (market) complete, and therefore a payoff can be dynamically replicated under this model. Indeed, if we take for granted that the dual is attained, the duality result implies that the payoff can be perfectly dynamically replicated

$$u_1(s_1) + u_2(s_2) + h(s_1)(s_2 - s_1) = c(s_1, s_2), \quad \mathbb{P}^* - \text{as}$$

for the optimal martingale measure \mathbb{P}^* .

Assum($\mathbb{P}^1, \mathbb{P}^2$): For the sake of simplicity, we will assume that $\delta F \equiv F_2 - F_1$ has a global unique maximum m . The general case is considered in [S4]. We define $\mathbb{P}^* \in \mathcal{M}(\mathbb{P}^1, \mathbb{P}^2)$ as

$$\begin{aligned} \mathbb{P}^*(ds_1, ds_2) &= \mathbb{P}^1(ds_1) (q(s_1)\delta_{T_u(s_1)}(ds_2) + (1 - q(s_1))\delta_{T_d(s_1)}(ds_2)), \\ q(x) &= \frac{x - T_d(x)}{T_u(x) - T_d(x)} \end{aligned}$$

with the maps $T_d(x) \leq x \leq T_u(x)$, T_u increasing, T_d decreasing defined by

$$\begin{aligned} T_u(x) &= T_d(x) = x, \quad x \leq m \\ T_u(x) &= F_2^{-1}(F_1(x) + \delta F(T_d(x))) \\ T_d'(x) &= -\frac{T_u(x) - x}{T_u(x) - T_d(x)} \frac{F_1'(x)}{F_2'(T_d(x)) - F_1'(T_d(x))} \end{aligned} \tag{1.10}$$

In Figure (1.2), we have plotted the maps T_d and T_u corresponding to two log-normal densities with variances 0.04 and 0.32 (increasing in the convex order). Note that the expression for the map $T_u(x)$ looks like the Fréchet-Hoeffding solution, except the presence of the additional term $\delta F(T_d(x))$ arising from our martingale condition.

We next introduce a remarkable triple of dual variables (u_1^*, u_2^*, h^*) corresponding to a smooth coupling function c . The dynamic hedging component h^* is defined up to an arbitrary constant by:

$$h^{*'} = \frac{c_{s_1}(\cdot, T_u) - c_{s_1}(\cdot, T_d)}{T_u - T_d}, \quad \forall x \geq m \tag{1.11}$$

The payoff function u_2^* is defined up to an arbitrary constant by:

$$u_2^{*'} = c_y(T_u^{-1}, \cdot) - h^* \circ T_u^{-1}, \quad \forall x \geq m \tag{1.12}$$

$$= c_y(T_d^{-1}, \cdot) - h^* \circ T_d^{-1}, \quad \forall x < m. \tag{1.13}$$

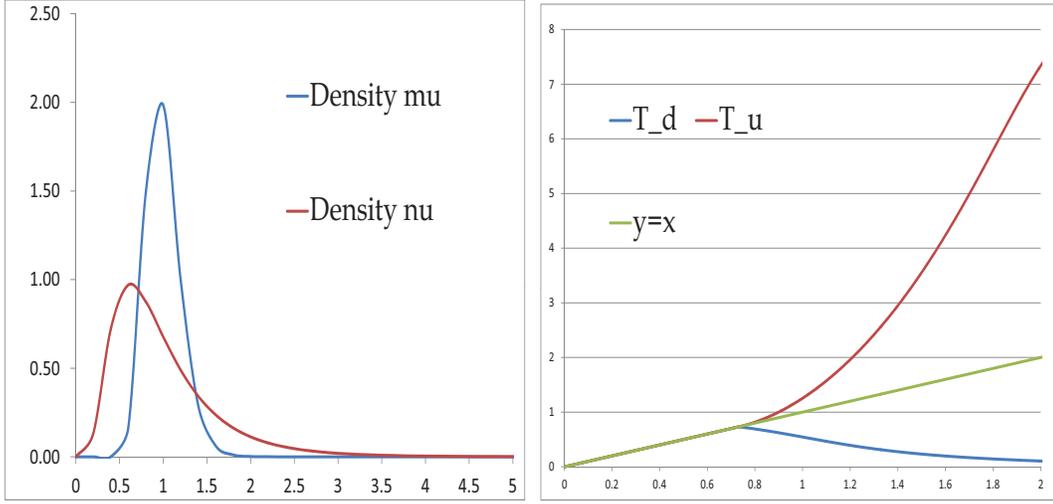


Figure 1.2: Maps T_d and T_u built from two log-normal densities μ and ν with variances 0.04 and 0.32. $m = 0.731$.

The corresponding function u_1^* is given by:

$$\begin{aligned} u_1^*(s_1) &= \mathbb{E}^{\mathbb{P}^*} [c(S_1, S_2) - u_2^*(S_2) | S_1 = s_1] \\ &= q(s_1)(c(s_1, \cdot) - u_2^*) \circ T_u(s_1) + (1 - q(s_1))(c(s_1, \cdot) - u_2^*) \circ T_d(s_1), \quad s_1 \in \mathbb{R}_+. \end{aligned} \quad (1.14)$$

Theorem 1.3.4 (see S4) Assume that $\mathbb{P}^1 < \mathbb{P}^2$ are Borel probability measures on \mathbb{R}_+ with first moments S_0 and assumption **Assum**($\mathbb{P}^1, \mathbb{P}^2$) holds. Assume further that $u_1^{*+} \in L^1(\mathbb{P}^1)$, $u_2^{*+} \in L^1(\mathbb{P}^2)$, and that the partial derivative of the coupling function $c_{s_1 s_2 s_2}$ exists and $c_{s_1 s_2 s_2} > 0$ on \mathbb{R}_+^2 . Then:

- (i) $(u_1^*, u_2^*, h^*) \in \mathcal{M}^*(\mathbb{P}^1, \mathbb{P}^2)$,
- (ii) the strong duality holds for the martingale transportation problem, \mathbb{P}^* is a solution of $\widetilde{\text{MK}}_2^*$, and (u_1^*, u_2^*, h^*) is a solution of $\widetilde{\text{MK}}_2$:

$$\mathbb{E}^{\mathbb{P}^*} [c(S_1, S_2)] = \widetilde{\text{MK}}_2^* = \widetilde{\text{MK}}_2 = \mathbb{E}^{\mathbb{P}^1} [u_1^*(S_1)] + \mathbb{E}^{\mathbb{P}^2} [u_2^*(S_2)].$$

In Table (1.1), we compare various results in optimal transport and in the martingale counterpart. We have seen that OT is linked to the Hamilton-Jacobi equation. In MOT (in particular in our approach to the Skorohod embedding problem), Hamilton-Jacobi-Bellman PDEs will naturally appear.

OT, MK_2	MOT, $\widetilde{\text{MK}}_2$
$\sup_{\mathbb{P}, S_1 \sim \mathbb{P}^1, S_2 \sim \mathbb{P}^2} \mathbb{E}^{\mathbb{P}}[c(S_1, S_2)]$	$\sup_{\mathbb{P}, S_1 \sim \mathbb{P}^1, S_2 \sim \mathbb{P}^2, \mathbb{E}[S_2 S_1]=S_1} \mathbb{E}^{\mathbb{P}}[c(S_1, S_2)]$
$\inf_{u_1, u_2} \mathbb{E}^{\mathbb{P}^1}[u_1(S_1)] + \mathbb{E}^{\mathbb{P}^2}[u_2(S_2)]$ $u_1(s_1) + u_2(s_2) \geq c(s_1, s_2)$	$\inf_{u_1, u_2, h} \mathbb{E}^{\mathbb{P}^1}[u_1(S_1)] + \mathbb{E}^{\mathbb{P}^2}[u_2(S_2)]$ $u_1(s_1) + u_2(s_2) + h(s_1)(s_2 - s_1) \geq c(s_1, s_2)$
$\inf_u \mathbb{E}^{\mathbb{P}^1}[u^*(S_1)] + \mathbb{E}^{\mathbb{P}^2}[u(S_2)]$	$\inf_u \mathbb{E}^{\mathbb{P}^1}[(c(S_1, \cdot) - u(\cdot))^{**}(S_1)] + \mathbb{E}^{\mathbb{P}^2}[u(S_2)]$
(co)-monotone map: $c_{12} > 0$	(co)-monotone martingale map: $c_{122} > 0$

Table 1.1: Optimal Transport versus Martingale Optimal Transport. $u^*(s_1) \equiv \sup_{s_2} \{c(s_1, s_2) - u(s_2)\}$ and $u^{**}(\cdot)$ denotes the concave envelope of $u(\cdot)$ which is the smallest concave function greater or equal to u .

1.3.3 Martingale Brenier's solution ♣

We consider MOT in \mathbb{R}_+^d and note S_1^i the i -component of $S_1 \in \mathbb{R}_+^d$. Note that this situation is not very common in finance as the (known) marginals are usually one-dimensional (e.g. Vanillas). A notable exception arise in fixed income and foreign exchange markets (see example 1.2.1) where Vanillas on spread swap rates, i.e., $S_1^2 - K S_1^1$, are quoted on the market. The dual reads

$$\widetilde{\text{MK}}_2 = \inf_{u_1 \in L^1(\mathbb{P}^1), u_2 \in L^1(\mathbb{P}^2), (h^i(\cdot))_{1 \leq i \leq d}} \mathbb{E}^{\mathbb{P}^1}[u_1(S_1)] + \mathbb{E}^{\mathbb{P}^2}[u_2(S_2)]$$

such that $u_1(s_1) + u_2(s_2) + \sum_{i=1}^d h^i(s_1)(s_2^i - s_1^i) \geq c(s_1, s_2)$, $\forall s_1, s_2 \in \mathbb{R}_+^d$. Taking for granted that the dual is attained, the (strong) duality result implies that

$$u_1(s_1) + u_2(s_2) + \sum_{i=1}^d h^i(s_1)(s_2^i - s_1^i) = c(s_1, s_2), \quad \mathbb{P}^* - \text{a.s.}$$

We have $d + 2$ unknown functions $(u_1, u_2, (h^i(\cdot))_{1 \leq i \leq d})$ (defined on (a subset of) \mathbb{R}_+^d) and it is tempting to guess that the optimal martingale measure \mathbb{P}^* is localized on some maps (T^α). For each map - denoted schematically by T - we should have:

$$u_1(s_1) + u_2(T(s_1)) + \sum_{i=1}^d h^i(s_1^i)(T^i(s_1) - s_1^i) = c(s_1, T(s_1)), \quad \forall s_1 \in \mathbb{R}^d \quad (1.15)$$

$$\partial_{s_2^i} u_2(T(s_1)) + h^i(s_1) = \partial_{s_2^i} c(s_1, T(s_1)), \quad \forall i = 1, \dots, d, \forall s_1 \in \mathbb{R}^d \quad (1.16)$$

On the primal side, the minimal number of maps (in order to get a martingale measure) should be $d + 1$:

$$\mathbb{P}^*(ds_1, s_2) = \sum_{i=1}^{d+1} q_i(s_1) \delta_{T_i(s_1)}(ds_2)$$

where the functions $(q_i)_{1 \leq i \leq d+1}$ are uniquely fixed by the algebraic equations:

$$\sum_{i=1}^{d+1} q_i = 1, \quad \sum_{i=1}^{d+1} q_i(s_1) (T_i(s_1) - s_1) = 0$$

An open problem is to extend our approach in $d = 1$ to \mathbb{R}_+^d . We intend to address this problem in our future research.

1.3.4 Symmetries in MOT [S4]

Martingale Spence-Mirrlees condition

The martingale counterpart of the Spence-Mirrlees condition is $c_{s_1 s_2 s_2} > 0$. This condition is natural in the present setting. Indeed, the optimization problem should not be affected by the modification of the coupling function from c to $\bar{c}(s_1, s_2) \equiv c(s_1, s_2) + U^1(s_1) + U^2(s_2) + h(s_1)(s_2 - s_1)$ for any $U^1 \in L^1(\mathbb{P}^1)$, $U^2 \in L^1(\mathbb{P}^2)$, and $h \in L^0$, the new optimal cost being $\widetilde{\text{MK}}_2 + \mathbb{E}^{\mathbb{P}^1}[U_1(S_1)] + \mathbb{E}^{\mathbb{P}^2}[U_2(S_2)]$. Since $\bar{c}_{s_1 s_2 s_2} = c_{s_1 s_2 s_2}$, it follows that the condition $c_{s_1 s_2 s_2} > 0$ is stable for the above transformation of the coupling function. In particular, note that $\bar{c}_{s_1 s_2} = c_{s_1 s_2} + h'(s_1) \neq c_{s_1 s_2}$.

Mirror coupling: the right-monotone martingale transport plan

Suppose that $c_{s_1 s_2 s_2} < 0$. Then, the upper bound $\widetilde{\text{MK}}_2$ is attained by the right-monotone martingale transport map

$$\mathbb{P}_*(ds_1, ds_2) = \mathbb{P}^1(ds_1) (q(s_1)\delta_{\bar{T}_u(s_1)}(ds_2) + (1 - q(s_1))\delta_{\bar{T}_d(s_1)}(ds_2)), \quad q(x) = \frac{x - \bar{T}_d(x)}{\bar{T}_u(x) - \bar{T}_d(x)}$$

where (\bar{T}_d, \bar{T}_u) is defined as in (1.10) with the pair of probability measures $(\bar{\mathbb{P}}^1, \bar{\mathbb{P}}^2)$:

$$F_{\bar{\mathbb{P}}^1}(s_1) \equiv 1 - F_{\mathbb{P}^1}(-s_1), \quad \text{and} \quad F_{\bar{\mathbb{P}}^2}(s_2) \equiv 1 - F_{\mathbb{P}^2}(-s_2).$$

To see this, we rewrite the optimal transportation problem equivalently with modified inputs:

$$\bar{c}(s_1, s_2) \equiv c(-s_1, -s_2), \quad \bar{\mathbb{P}}^1((-\infty, s_1]) \equiv \mathbb{P}^1([-s_1, \infty)), \quad \bar{\mathbb{P}}^2((-\infty, s_2]) \equiv \mathbb{P}^2([-s_2, \infty)),$$

so that $\bar{c}_{s_1 s_2 s_2} > 0$, as required in Theorem 1.3.4. Note that the martingale constraint is preserved by the map $(s_1, s_2) \rightarrow (-s_1, -s_2)$.

Suppose that $c_{s_1 s_2 s_2} > 0$. Then, the lower bound problem is explicitly solved by the right-monotone martingale transport plan. Indeed, it follows from the first part of the present remark that:

$$\inf_{\mathbb{P} \in \mathcal{M}(\mathbb{P}^1, \mathbb{P}^2)} \mathbb{E}^{\mathbb{P}}[c(S_1, S_1)] = - \sup_{\mathbb{P} \in \mathcal{M}(\mathbb{P}^1, \mathbb{P}^2)} \mathbb{E}^{\mathbb{P}}[-c(S_1, S_2)] = \mathbb{E}^{\bar{\mathbb{P}}^*}[c(S_1, S_2)]$$

1.3.5 Martingale McCann's interpolation ♣

In this section, we recall McCann's interpolation theory. Our reminder follows closely Chapter 5 in Villani's book [107]. Here we consider the case \mathbb{R}^d , although our martingale extension focuses only on the real line. Two probability measures \mathbb{P}_0 and \mathbb{P}_1 (not necessarily in convex order here) can be (trivially) linearly interpolated by $\mathbb{P}_t = (1 - t)\mathbb{P}_0 + t\mathbb{P}_1$ with $t \in [0, 1]$. In mathematical

physics, one often consider minimization of functionals on the space of probability measures $P(\mathbb{R}^d)$: $\inf_{\mathbb{P} \in P(\mathbb{R}^d)} \mathcal{F}(\mathbb{P})$. The proof of the existence of a unique minimizer is greatly simplified if we can prove that the functional \mathcal{F} is strictly convex. However classical examples such that

$$\mathcal{F}_2(\mathbb{P}) \equiv \int_{\mathbb{R}^d \times \mathbb{R}^d} W(x-y) \mathbb{P}(dx) \mathbb{P}(dy) \quad (1.17)$$

do not satisfy this property even if W is strictly convex. Indeed, a straightforward computation gives

$$\frac{d^2}{dt^2} \mathcal{F}_2(\mathbb{P}_t) = 2 \int W(x-y) (\mathbb{P}_0(dx) \mathbb{P}_0(dy) - \mathbb{P}_1(dx) \mathbb{P}_0(dy) - \mathbb{P}_0(dx) \mathbb{P}_1(dy) + \mathbb{P}_1(dx) \mathbb{P}_1(dy))$$

for which it is not possible to conclude the convex property. This was McCann's original motivation [84] for introducing a new notion of probability interpolation and convexity which can handle the above example. His approach is strongly linked to Brenier's theorem in optimal transport as described below.

Let's take a strictly concave cost $c = c(s_1 - s_2)$ on \mathbb{R}^d . According to Brenier's theorem, there exists a gradient of a convex function u such that

$$[1 - \nabla c^*(\nabla u)] \# \mathbb{P}_0 = \mathbb{P}_1$$

McCann's displacement interpolation $[\mathbb{P}_0, \mathbb{P}_1]_t$ of two probability measures \mathbb{P}_0 and \mathbb{P}_1 is then defined by

$$[\mathbb{P}_0, \mathbb{P}_1]_t \equiv [1 - t \nabla c^*(\nabla u)] \# \mathbb{P}_0, \quad t \in [0, 1]$$

Note that by construction $\mathbb{P}_{t=0} = \mathbb{P}_0$ and $\mathbb{P}_{t=1} = \mathbb{P}_1$. A subset \mathcal{P} of $P(\mathbb{R}^d)$ is said to be displacement convex if it is stable under displacement interpolation. A functional \mathcal{F} defined on a displacement convex subset \mathcal{P} is displacement (resp. strictly) convex if for all $\mathbb{P}_0, \mathbb{P}_1$ in \mathcal{P} , the function $t \rightarrow \mathcal{F}([\mathbb{P}_0, \mathbb{P}_1]_t)$ is (resp. strictly) convex. In particular, the functional \mathcal{F}_2 is displacement convex if W is convex [84]. One can then show the existence of a unique minimizer [84] of \mathcal{F} using this displacement convex property. For applications in mathematical finance, one can consider minimizations of functionals defined on a subset of $P(\mathbb{R})$ totally ordered with respect to the convex order. Note that in comparison with the usual convex interpolation, McCann's interpolation does not preserve the convex order property: If $\mathbb{P}_0 \leq \mathbb{P}_1$, then we have *not* $\mathbb{P}_0 \leq [\mathbb{P}_0, \mathbb{P}_1]_t \leq \mathbb{P}_1$.

Using the framework of MOT, we define the convex order interpolation of two measures $\mathbb{P}_0 \leq \mathbb{P}_1$ defined on \mathbb{R} as

Definition 1.3.5 (Martingale convex interpolation)

$$[\mathbb{P}_0, \mathbb{P}_1]_t = \text{Law}(X_t), \quad t \in [0, 1]$$

where the random variable X_t is defined by a two step Markov chain: $X_t = X(1-t) + tT_u(X)$ with probability $q(x) \equiv \frac{x - T_d(x)}{T_u(x) - T_d(x)}$, $X_t = X(1-t) + tT_d(X)$ with probability $1 - q(x)$ and $X \sim \mathbb{P}_0$.

By construction, this interpolation preserves the convex order property as this Markov chain is a martingale. Note that the linear interpolation $\mathbb{P}_t = (1-t)\mathbb{P}_0 + t\mathbb{P}_1$ preserves also the convex order property. We explain in the next section why our interpolation seems better.

Application

We introduce a subset $\mathcal{P}_{\text{c.o.}}$ of the space of probability measures, absolutely continuous with respect to the Lebesgue measure on \mathbb{R} , totally ordered with respect to the convex order property. Here we assume that this space is stable under convex order displacement interpolation: if \mathbb{P}_0 and \mathbb{P}_1 are in $\mathcal{P}_{\text{c.o.}}$ (and therefore without loss of generality we take $\mathbb{P}_0 \leq \mathbb{P}_1$) then $[\mathbb{P}_0, \mathbb{P}_1]_t \in \mathcal{P}$ for $t \in [0, 1]$. Note that \mathcal{P} is a convex set in the usual sense. We say that a functional on $\mathcal{P}_{\text{c.o.}}$ is displacement (resp. strictly) convex order if the map $t \rightarrow F(\mathbb{P}_t)$ is (resp. strictly) convex on $[0, 1]$.

The functional \mathcal{F}_2 introduced above (1.17) is an example of displacement (resp. strictly) convex order functional. This example is not convex in the usual sense. This is why our notion seems better. Our definition (in particular the fact that q remains time independent under convex order interpolation) seems to be the right definition. As an application of the convex order interpolation, we have

Theorem 1.3.6 (see a similar theorem in OT, Theorem 5.32 in [107]) *Consider the following functional $\mathcal{F}_2(\mathbb{P})$ defined on $\mathcal{P}_{\text{c.o.}}$ (see the definition above) where W is strictly convex. Then, there is at most one minimizer for $\mathcal{F}_2(\mathbb{P})$ on $\mathcal{P}_{\text{c.o.}}$.*

This result is obvious as $\mathcal{F}_2(\mathbb{P})$ is strictly convex with respect to the martingale convex interpolation.

1.3.6 Multi-marginals extension [S4]

A natural extension of our 2-period MOT is to consider the multi-marginal case ($n > 2$) where the dual formulation is

$$\widetilde{\text{MK}}_n = \sup_{\mathbb{P} \in \mathcal{M}(\mathbb{P}^1, \dots, \mathbb{P}^n)} \mathbb{E}^{\mathbb{P}}[c(S_1, \dots, S_n)]$$

where $\mathcal{M}(\mathbb{P}^1, \dots, \mathbb{P}^n) = \{\mathbb{P} : S_i \stackrel{\mathbb{P}}{\sim} \mathbb{P}^i, \mathbb{E}_{i_{i-1}}^{\mathbb{P}}[S_i] = S_{i-1}, i = 1, \dots, n\}$. This problem was solved in the classical optimal transport problem in [29, 54] (see Section 1.2.4). In our present martingale version, by using a Markov property, and a specific class of cost functions, $c = \sum_{i=1}^{n-1} c^i(s_i, s_{i+1})$ with $c_{s_i s_{i+1} s_{i+1}}^i > 0$, it is natural to guess that the optimal martingale measure will be supported by $2 \times n$ maps (T_d^i, T_u^i) satisfying ODEs (1.10) with F_1 and F_2 replaced by F_i and F_{i+1} . More precisely, we have

Theorem 1.3.7 (see S4) *Suppose $\mathbb{P}^1 \leq \dots \leq \mathbb{P}^n$ in convex order, with finite first moment, $\mathbb{P}^1, \dots, \mathbb{P}^{n-1}$ have no atoms, and let Assumption **Assum**($\mathbb{P}^i, \mathbb{P}^{i+1}$) hold true for $\delta F = F_{i+1} - F_i$, for all $1 \leq i < n$. Assume further that*

- c^i have linear growth, that the cross derivatives c_{xy}^i exist and satisfy $c_{xy}^i > 0$,
- u_i^{1*}, u_i^{2*} satisfy the integrability conditions $(u_i^{1*})^+ \in L^1(\mathbb{P}^i)$, $(u_i^{2*})^+ \in L^1(\mathbb{P}^{i+1})$. u_i^{1*}, u_i^{2*} are defined as in (1.11, 1.12, 1.14) with c^i substituted to c and (T_u^i, T_d^i) substituted to (T_u, T_d) .

Maturity (years)	VS _{mkt}	Upper	Lower
0.4	18.47	18.45	16.73
0.6	19.14	18.70	17.23
0.9	20.03	19.63	17.89
1.4	21.77	21.62	19.03
1.9	22.89	23.06	19.63

Table 1.2: Implied volatility for discrete-monitored variance swap as a function of the maturity. Lower/upper bounds versus market prices (quoted in volatility $\times 100$).

Then, the strong duality holds, the transference map

$$\mathbb{P}^*(ds) = \mathbb{P}^1(ds_1) \prod_{i=1}^{n-1} \left(q_i(s_i) \delta_{T_u^i(s_i)}(ds_{i+1}) + (1 - q_i(s_i)) \delta_{T_d^i(s_i)}(ds_{i+1}) \right)$$

is optimal for the martingale transportation problem $\widetilde{\text{MK}}_n$, and (u^*, h^*) is optimal for the dual problem, i.e.,

$$\mathbb{E}^{\mathbb{P}^*} [c(S_1, \dots, S_n)] = \sum_{i=1}^n \mathbb{E}^{\mathbb{P}^i} [u_i^*].$$

with $u_i^*(s_i) \equiv 1_{i < n} u_i^{1*} + 1_{i > 1} u_{i-1}^{2*}$, $i = 1, \dots, n$.

As an useful application in mathematical finance, this result gives the robust lower and upper bounds of a discrete-monitored variance swap delivering at a maturity T , the discrete realized variance of log-returns: $-2/T \sum_{i=1}^N \ln^2 s_{i+1}/s_i$. We have compared these bounds against market values denoted VS_{mkt} for the DAX index (2-Feb-2013) and different maturities (see Table 1.2). The prices are quoted in volatility $\times 100$. Note that for maturities less than one year, our upper bound is below the market price, highlighting an arbitrage opportunity. In practice, this arbitrage disappears if we include transaction costs for trading vanilla options with low/high strikes. Moreover, we have assumed that vanilla options with all maturities are traded.

1.3.7 Infinitely-many marginals extension [S3]

At this point, we now introduce a continuous time martingale transportation (MT) problem with full marginals constraints, as the limit of the multi-marginals MT recalled in previous section. Namely, given a family of marginals $(\mathbb{P}^t)_{t \in (0, T]}$, we consider all the continuous time martingales satisfying the marginal constraints, and optimize w.r.t. a class of cost functions.

Let $\Omega \equiv D([0, T], \mathbb{R}_+)$ denote the canonical space of all càdlàg paths on $[0, T]$, S the canonical process and $\mathcal{F} = (\mathcal{F}_t)_{0 \leq t \leq T}$ the canonical filtration generated by S . We denote by \mathcal{M}_∞ the collection of all martingale measures on Ω , i.e. the collection of all probability measures on Ω under which the canonical process S is a martingale. $S_{t=0}$ is fixed to S_0 . By Karandikar [82],

there is a non-decreasing process $(\langle S \rangle_t)_{t \in [0,1]}$ defined on Ω which coincides with the \mathbb{P} -quadratic variation of S , \mathbb{P} -a.s. for every martingale measure $\mathbb{P} \in \mathcal{M}_\infty$. Denote also by $\langle S \rangle^c$ the continuous part of the non-decreasing process $\langle S \rangle$.

Given a family of marginals $(\mathbb{P}^t)_{0 < t \leq T}$, denote by $\mathcal{M}_\infty((\mathbb{P}^t)_{0 < t \leq T})$ the collection of all martingale measures on Ω such that $S_t \stackrel{\mathbb{P}}{\sim} \mathbb{P}^t$ for all $t \in (0, T]$. In particular, following Kellerer [83], $\mathcal{M}_\infty((\mathbb{P}^t)_{0 < t \leq T})$ is nonempty if and only if the family $(\mathbb{P}^t)_{0 < t \leq T}$ is non-decreasing in convex ordering and $t \mapsto \mathbb{P}^t$ is right-continuous.

We introduce a cost function $C : \Omega \rightarrow \mathbb{R} \cup \{-\infty, \infty\}$ by

$$C(\mathbf{x}) \equiv \frac{1}{2} \int_0^T c_{yy}(\mathbf{x}_t, \mathbf{x}_t) d\langle \mathbf{x} \rangle_t^c + \sum_{0 \leq t \leq T} c(\mathbf{x}_{t-}, \mathbf{x}_t), \quad \forall \mathbf{x} \in \Omega,$$

where the integral and the sum are defined as the difference of the positive and negative parts, under the convention $\infty - \infty = -\infty$, e.g.

$$\int_0^1 c_{yy}(\mathbf{x}_t, \mathbf{x}_t) d\langle \mathbf{x} \rangle_t^c \equiv \int_0^1 (c_{yy}(\mathbf{x}_t, \mathbf{x}_t))^+ d\langle \mathbf{x} \rangle_t^c - \int_0^1 (c_{yy}(\mathbf{x}_t, \mathbf{x}_t))^- d\langle \mathbf{x} \rangle_t^c$$

The cost function c is assumed to satisfy

Assumption .1 *The cost function $c : \mathbb{R}_+^2 \rightarrow \mathbb{R}$ is in $C^3(\mathbb{R}_+^2)$ and satisfies*

$$c(x, x) = c_y(x, x) = 0 \quad \text{and} \quad c_{xyy}(x, y) > 0, \quad \forall x \leq y \quad (1.18)$$

We then formulate a continuous time MT problem under full marginals constraints by

$$\widetilde{\text{MK}}_\infty \equiv \sup_{\mathbb{P} \in \mathcal{M}_\infty((\mathbb{P}^t)_{0 < t \leq T})} \mathbb{E}^\mathbb{P}[C(S.)]. \quad (1.19)$$

Martingales matching marginals: An optimal transport point of view

A Monge-Kantorovich dual formulation for $\widetilde{\text{MK}}_\infty$ is provided in [S3]. The solution of this MOT should provide a martingale measure with marginals $(\mathbb{P}^t)_{0 < t \leq T}$. Building such a martingale measure is a difficult task and, although only few examples are known, this issue has generated a lot of research in probability and mathematical finance. In mathematical finance layman's terms, this measure is connected to an arbitrage-free model calibrated exactly to the full Vanilla smile. The first example, commonly used by practitioners in quantitative finance in part due to this property, is the local volatility model introduced by B. Dupire [41] and defined by a one-dimensional Markovian SDE

$$dS_t = \sigma_{\text{loc}}(t, S_t) dW_t, \quad \sigma_{\text{loc}}(t, K)^2 = 2 \frac{\partial_t C(t, K)}{\partial_K^2 C(t, K)}$$

with $C(t, K) = \mathbb{E}^{\mathbb{P}^t}[(S_t - K)^+]$. A second example is provided by so-called local stochastic volatility models that are described by nonlinear McKean SDEs:

$$dS_t = a_t \frac{\sigma_{\text{loc}}(t, S_t)}{\sqrt{\mathbb{E}[a_t^2 | S_t]}} dW_t$$

We will expand on this model in the last chapter when we discuss the particle method - a subject connected to classical optimal transport. The last example pops up when considering the Skorohod embedding problem (in short SEP) - see also next section: It consists in finding a stopping time τ such that $B_\tau \sim \mathbb{P}^T$ with B a Brownian motion, \mathbb{P}^T a marginal distribution and $B_{t \wedge \tau}$ is required to be uniformly integrable. From the Dambis-Dubins-Schwarz theorem, all continuous local martingale S_t (with $\lim_{t \rightarrow \infty} \langle S \rangle_t = \infty$) can be converted into a Brownian subordinated by its quadratic variation: $S_t = B_{\langle S \rangle_t}$. Taking $\langle S \rangle_T = \tau$, S_t describes an arbitrage-free model calibrated to T -Vanillas. A solution of SEP is the Azema-Yor solution [11] where the stopping time τ is described at the first time to reach a barrier:

$$\tau = \inf\{t > 0 : \Psi_{\text{AY}, \mathbb{P}^T}(B_t) \leq B_t^*\}$$

with $B_t^* = \max_{0 < s < t} B_s$ the running maximum and Ψ_{AY} is the Hardy-Littlewood barycenter function depending on the marginal \mathbb{P}^T (implied from T -Vanilla options)

$$\Psi_{\text{AY}, \mathbb{P}^T}(x) = \frac{\mathbb{E}^{\mathbb{P}^T}[S_T 1_{S_T > x}]}{\mathbb{E}^{\mathbb{P}^T}[1_{S_T > x}]}$$

From this solution, one can derive a martingale in $\mathcal{M}_\infty((\mathbb{P}^t)_{0 < t \leq T})$ under the strong assumption: the barycenter functions $\Psi_{\text{AY}, \mathbb{P}^t}$ are increasing in time [90], i.e. $\Psi_{\text{AY}, \mathbb{P}^s}(x) \leq \Psi_{\text{AY}, \mathbb{P}^t}(x)$ for all $0 < s < t$. The corresponding martingale is $S_t = B_{\tau_t}$ with

$$\tau_t = \inf\{s > 0 : \Psi_{\text{AY}, \mathbb{P}^t}(B_s) \leq B_s^*\} \quad (1.20)$$

This martingale is a pure jump Lévy process for which one can explicitly characterize its infinitesimal generator [90]. Alternative pure jump Lévy processes with a local intensity, depending on the time t and S_t , have been built in [31] and can be calibrated to all continuous-time marginal distributions $(\mathbb{P}^t)_{0 < t \leq T}$ totally ordered with respect to the convex order. In the next section, we show that the martingales obtained from the limit $n \rightarrow \infty$ of our discrete optimal measures \mathbb{P}^* and \mathbb{P}_* belong to the class of local Lévy processes (see [S3] for technical details). These martingales are remarkable as they are the optimal solutions (upper and lower bounds) of our continuous-time martingale optimal transport problem $\bar{\text{MK}}_\infty$ (under assumption .1).

Infinitely-many marginals: optimal solutions [S3]

We denote $f(t, \cdot)$ (resp. $F(t, \cdot)$) the continuous strictly positive density (resp. cumulative distribution) of S_t at time t implied from t -Vanilla option prices. We will assume that there exists a unique local maximizer of $\partial_t F(t, \cdot)$ that we denote $m(t)$ (see [S3] for the general case).

Heuristic derivation

In the continuous-time limit $\Delta t \equiv t_2 - t_1 \rightarrow 0$, one can guess that the solution of ODEs (1.10) can be written at the first order in Δt for $s > m(t)$ as

$$\begin{aligned} T_u(t, s) &= s + j_u(t, s)\Delta t \\ T_d(t, s) &= s - j_d(t, s)\Delta t \end{aligned}$$

with $j_d(t, s) = j_u(t, s) = 0$, $\forall s \leq m(t)$. Plugging this expression into the ODE system, we get for all $s > m(t)$:

$$1 - \partial_s j_d(t, s) = -\frac{j_u(t, s)}{j_d(t, s)} \frac{f(t, s)}{\partial_t f(t, s - j_d(t, s))} \quad (1.21)$$

$$\partial_s(j_u(t, s)f(t, s)) + \partial_t f(t, s) = -\frac{j_u(t, s)}{j_d(t, s)} f(t, s) \quad (1.22)$$

The master equation, associated to this binomial model, is

$$\begin{aligned} f(t + \Delta t, s) &= \left(f(t, s) \mathbf{1}_{s \leq m(t)} - \frac{j_u(t, \underline{s})}{j_d(t, \underline{s})} \mathbf{1}_{s < m(t)} \frac{f(t, \underline{s})}{(1 - \partial_s j_d(t, \underline{s}))} \Delta t \right) \\ &+ \left(1 - \frac{j_u(t, \bar{s})}{j_d(t, \bar{s})} \Delta t \right) \frac{f(t, \bar{s})}{1 + \partial_s j_u(t, \bar{s}) \Delta t} \mathbf{1}_{s > m(t)} \end{aligned}$$

with $\underline{s} - j_d(t, \underline{s}) = s$, $\bar{s} + j_u(t, \bar{s}) \Delta t = s$, and where we have used that the probability to jump downwards at t is $q(t, s) = \frac{j_u(t, s)}{j_d(t, s)} \Delta t$. By taking formally the limit $\Delta t \rightarrow 0$ in the above equation, we obtain the following Fokker-Planck equation:

$$\partial_t f(t, s) = -\frac{j_u(t, \underline{s})}{j_d(t, \underline{s})} \frac{f(t, \underline{s})}{(1 - j'_d(t, \underline{s}))} \mathbf{1}_{s < m(t)} - \frac{j_u(t, s)}{j_d(t, s)} f(t, s) \mathbf{1}_{s > m(t)} - \partial_s(j_u(t, s)f(t, s)) \mathbf{1}_{s > m(t)} \quad (1.23)$$

The infinitesimal generator appearing in the above Fokker-Planck equation is obtained from a compensated pure jump (Markovian) martingale with compensator

$$\nu(dt, dz) = \delta(z + j_d(t, S_{t-})) \frac{j_u(t, S_{t-})}{j_d(t, S_{t-})} \mathbf{1}_{S_{t-} > m(t)} dz dt \quad (1.24)$$

The dynamics for the stock price is given by

$$dS_t = \int_{\mathbb{R}} z (N(dt, dz) - \nu(dt, dz))$$

where $N(dt, dz)$ is the counting measure associated with the (downwards) jumps. This jump process corresponds to a local Lévy model as introduced in [29]. A similar construction can be obtained for the anti-monotone rearrangement martingale which corresponds to compensated upward jumps. We obtain :

Theorem 1.3.8 (see S3) *The martingale transport problem (1.19) is solved by the local Lévy process (1.24). Moreover, we have*

$$\mathbb{E}^{\mathbb{P}^*} [C(X.)] = \mu(\lambda^*),$$

where the optimal value is given by

$$\mu(\lambda^*) = \int_0^T \int_{m(t)}^{\infty} \frac{j_u(t, s)}{j_d(t, s)} c(s, s - j_d(t, s)) f(t, s) ds dt.$$

Fake Brownian motion [S3]

These two jump processes give rise two new examples of (non-continuous) fake Brownian motions (if we take for f_t the density of a Brownian motion). A fake Brownian motion is a local martingale with the same marginals as a Brownian motion. We also refer to Albin [4], Fan, Hamza and Klebaner [48], Hamza and Klebaner [60], Hirsch et al. [71], Hobson [72], Oleszkiewicz [95] etc. for other solutions and related results.

Let $\mathbb{P}^t \equiv \mathcal{N}(0, t)$. By direct computation, we have $m(t) = -\sqrt{t}$. In this case, it follows that $T_d(t, s)$ is defined by the equation:

$$\int_{T_d(t,s)}^s (s - \xi)(\xi^2 - t)e^{-\xi^2/2t} d\xi = 0 \quad \text{for all } s \geq m(t).$$

By direct change of variables, this provides the scaled solution $T_d(t, s) \equiv t^{1/2}\widehat{T}_d(t^{-1/2}s)$, where:

$$\widehat{T}_d(s) \leq -1 \quad \text{is defined for all } s \geq -1 \text{ by } \int_{\widehat{T}_d(s)}^s (s - \xi)(\xi^2 - 1)e^{-\xi^2/2} d\xi = 0.$$

i.e.

$$e^{-\widehat{T}_d(s)^2/2} \left(1 + \widehat{T}_d(s)^2 - s\widehat{T}_d(s)\right) = e^{-s^2/2}.$$

Similarly, we see that $j_u(t, s) \equiv t^{-1/2}\widehat{j}_u(t^{-1/2}s)$, where

$$\widehat{j}_u(s) \equiv \frac{1}{2} \left[s - \widehat{T}_d(s)e^{-(\widehat{T}_d(s)^2 - s^2)/2} \right] = \frac{1}{2} \left[s - \frac{\widehat{T}_d(s)}{1 + \widehat{T}_d(s)^2 - s\widehat{T}_d(s)} \right] \quad \text{for all } s \geq -1.$$

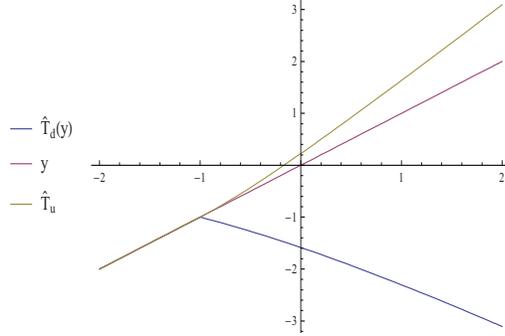
We have plotted the maps $\widehat{T}_d(s)$ and $\widehat{T}_u(s) \equiv s + \widehat{j}_u(s)$ in Figure 1.3. A similar construction can be achieved for self-similar marginals (see [S3]).

1.4 Constrained martingale optimal transport

Some variants in optimal transport have been studied recently. In [85], the authors introduce an optimal transportation with capacity constraints, which consists in minimizing a cost among joint densities \mathbb{P} with marginals \mathbb{P}^1 and \mathbb{P}^2 and under the *capacity* constraint

$$\mathbb{P}(s_1, s_2) \leq \bar{\mathbb{P}}(s_1, s_2)$$

for some prior joint density $\bar{\mathbb{P}}(s_1, s_2)$. In [29], optimal transports with congestion are considered. In this section, we present two variants of our martingale optimal transport. The first one involves a constraint on the VIX future [S1] and the second one involves a penalty entropy term [A2].

Figure 1.3: Fake Brownian motion: Maps \hat{T}_d and \hat{T}_u .

1.4.1 VIX constraints [S1]

VIX futures and VIX options, traded on the CBOE, have become popular volatility derivatives. The VIX index at a future expiry t_1 is by definition the 30-day variance swap volatility, computed by replication using market prices of listed S&P500 options (at t_1):

$$\begin{aligned} \text{VIX}_{t_1}^2 &\equiv -\frac{2}{\Delta} \mathbb{E}_{t_1} \left[\ln \left(\frac{S_{t_2}}{F_{t_1}^{t_2}} \right) \right] \\ &= \frac{2}{\Delta} \left(\int_0^{S_{t_1}} \frac{P(t_1, t_2, K)}{K^2} dK + \int_0^{F_{t_1}^{t_2}} \frac{C(t_1, t_2, K)}{K^2} dK \right) \end{aligned} \quad (1.25)$$

with $\Delta = t_2 - t_1$ and $P(t_1, t_2, K)$ (resp. $C(t_1, t_2, K)$) the undiscounted market price at t_1 of a put (resp. call) option with strike K and maturity t_2 . The payoff of a call option on VIX expiring at t_1 with strike K is $(\text{VIX}_{t_1} - K)^+$. Below, the market value (at $t = 0$) for the VIX future (i.e., $K = 0$) is denoted VIX.

For technical reason, we will assume that the random variables $(S_{t_1}, S_{t_2}, \text{VIX}_{t_1})$ are supported on a compact interval $I_1 \times I_2 \times I_X \subset (\mathbb{R}_+^*)^3$. For further reference, we denote by $\mathcal{M}(\mathbb{P}^1, \mathbb{P}^2, \text{VIX})$ the set of all martingale measures \mathbb{P} on the (pathspace) $I_1 \times I_2$ having marginals $\mathbb{P}^1, \mathbb{P}^2$ with mean S_0 and such that $\text{VIX} = \mathbb{E}^{\mathbb{P}}[\text{VIX}_{t_1}]$.

We define our constrained optimal transport for a VIX call option expiring at t_1 with strike K as

Definition 1.4.1

$$\text{MK}_{\text{vix}} \equiv \inf_{u_1 \in L^1(\mathbb{P}^1), u_2 \in L^1(\mathbb{P}^2), \lambda \in \mathbb{R}, h_S, h_X} \mathbb{E}^{\mathbb{P}^1}[u_1(S_{t_1})] + \mathbb{E}^{\mathbb{P}^2}[u_2(S_{t_2})] + \lambda \text{VIX}$$

such that for all $(s_1, s_2, x) \in I_1 \times I_2 \times I_X$,

$$u_1(s_1) + u_2(s_2) + \lambda\sqrt{x} + h_S(s_1, x)(s_2 - s_1) + h_X(s_1, x) \left(-\frac{2}{\Delta} \ln \left(\frac{s_2}{s_1} \right) - x \right) \geq (\sqrt{x} - K)^+ \quad (1.26)$$

where the functions $h_S, h_X : I_1 \times I_X \rightarrow \mathbb{R}$ are assumed to be bounded continuous functions on $I_1 \times I_X$, $u_1 \in L^1(\mathbb{P}^1)$ and $u_2 \in L^1(\mathbb{P}^2)$. Note that this defines a *linear* semi-infinite infinite-dimensional programming problem. The variable x should be interpreted as the t_1 -value of a log-contract $-2/\Delta \ln \frac{s_2}{s_1}$, i.e., the square of the VIX index $\text{VIX}_{t_1}^2$. Note that we have ensured that the above inequality is valid for all s_1 and s_2 , not just $\mathbb{P}^1 \times \mathbb{P}^2$ almost surely. As explained in [107] (see section 1.1), this can be achieved by allowing u_1 and u_2 to take values in $\mathbb{R} \cup \{+\infty\}$. Indeed, we introduce negligible sets N_1 and N_2 such that the inequality holds true for all $(s_1, s_2) \in N_1^c \times N_2^c$, and redefine the values of u_1 and u_2 to be $+\infty$ on N_1, N_2 respectively.

We establish a dual version which is connected to a martingale optimal transport problem:

Theorem 1.4.2 (duality, see S1) *Assume that $\mathbb{P}^1, \mathbb{P}^2$ are probability measures respectively on I_1 and I_2 such that $\mathcal{M}(\mathbb{P}^1, \mathbb{P}^2, \text{VIX})$ is non-empty. Then,*

$$\text{MK}_{\text{vix}} = \max_{\mathbb{P} \in \mathcal{M}(\mathbb{P}^1, \mathbb{P}^2, \text{VIX})} \mathbb{E}[(\text{VIX}_{t_1} - K)^+]$$

This is part of the proof that we have a max and not only a sup, meaning that the seller's price (resp. buyer's price) is attained by a martingale measure, i.e., a model, calibrated to the t_1 and t_2 Vanilla smiles and to the VIX future.

The dual corresponds to the maximization of the expectation of a VIX payoff with respect to a martingale measure with marginals $\mathbb{P}^1, \mathbb{P}^2$ and with the constraint on the VIX future $\text{VIX} = \mathbb{E}^{\mathbb{P}}[\sqrt{\mathbb{E}_{t_1}^{\mathbb{P}} - (2/\Delta) \ln S_{t_2}/S_{t_1}}]$. Note that as this additional constraint, not present in the original martingale optimal transport, is nonlinear with respect to the (martingale) measure \mathbb{P} , this optimal transport problem is more involved. Our variant involves a martingale constraint but also a *non linear* constraint $\text{VIX} = \mathbb{E}^{\mathbb{P}}[\sqrt{\mathbb{E}_{t_1}^{\mathbb{P}}[-(2/\Delta) \ln S_{t_2}/S_{t_1}]}]$. As a crucial step, by introducing a delta hedging on a forward log-contract $-2/\Delta \ln s_2/s_1$, this problem has been converted into a linear programming problem that can be solved with a simplex algorithm. Note that a similar trick is used in [25] for converting a quantile hedging approach into a super-replication problem. Under technical conditions on the marginal \mathbb{P}^2 , this optimal transport can be solved explicitly (see [S1]).

1.4.2 Entropy penalty [A2]

The main drawback of the (robust) super-replication approach, appearing in the definition of the martingale optimal transport, is that the martingale measure (i.e., arbitrage-free model), which achieves the super-replication strategy, can be very different from those generated by (stochastic volatility) diffusive models traders commonly use. As an example, the upper bound given marginals (such that the associated Hardy-Littlewood barycenter functions are increasing in time) for an increasing payoff on the maximum of a martingale is reached by Azema-Yor's martingale (see equation 1.20) belonging to the class of local jump Lévy models. In the following, we show that this drawback can be circumvented. To this end, we introduce the Kullback-Leibler relative entropy between two probability measures \mathbb{P} and \mathbb{P}^0 (see [107] for a review of various properties of this distance):

$$\begin{aligned} H(\mathbb{P}, \mathbb{P}^0) &= \mathbb{E}^{\mathbb{P}} \left[\ln \frac{d\mathbb{P}}{d\mathbb{P}^0} \right], \quad \mathbb{P} \text{ is absolutely continuous w.r.t. } \mathbb{P}^0 \\ &= +\infty, \text{ otherwise} \end{aligned}$$

Following the same route as in the super-replication section, we impose that our model matches market marginals at each date $(t_i)_{i=1, \dots, n}$ (eventually market instruments $(c_a)_{a=1, \dots, N}$)

$$\widetilde{\text{MK}}_n^\lambda \equiv \sup \{ \mathbb{E}^{\mathbb{P}} [c(S_{t_1}, \dots, S_{t_n})] : \mathbb{P} \in \mathcal{M}(\mathbb{P}^1, \dots, \mathbb{P}^n), H(\mathbb{P}, \mathbb{P}^0) \leq \lambda \} \quad (1.27)$$

We have added the constraint $H(\mathbb{P}, \mathbb{P}^0) \leq \lambda$ with λ a positive parameter. By setting $\lambda = 0$, the feasible set is empty except if the prior measure \mathbb{P}^0 is a martingale measure and satisfies the marginal constraints. In this case, we get

$$\widetilde{\text{MK}}_n^0 = \mathbb{E}^{\mathbb{P}^0} [c]$$

By taking the limit $\lambda = \infty$, we obtain our initial martingale optimal transport:

$$\widetilde{\text{MK}}s_n = \widetilde{\text{MK}}_n$$

The parameter λ provides an interpolation between prices given either by the prior (risk-neutral) model or the (robust) super-replication strategy.

Such an approach was explored in [7], [8], [9]. The martingality constraints seem to have been unnoticed, in particular in [7] and [9], resulting in pricing models that are not arbitrage-free. The authors stated in [7]: “In this paper, we study an algorithm for selecting a risk-neutral probability that minimizes the relative entropy, or Kullback-Leibler entropy, with respect to a given prior distribution.” We would like to emphasize that the density is not risk-neutral as long as the martingality constraint is not fulfilled. This is confirmed in proposition 4 in [8], where the drift of the diffusion measure \mathbb{P}^* is computed and found to be different from the risk-free interest rate.

The probability measure \mathbb{P} which achieves the maximum (the supremum is attained - see Theorem 1.4.3) is payoff-dependent. In order to get a payoff-independent measure, we could introduce a new primal problem:

$$\mathbb{P}^0 \equiv \inf \{ H(\mathbb{P}, \mathbb{P}^0) : \mathbb{P} \in \mathcal{M}(\mathbb{P}^1, \dots, \mathbb{P}^n) \}$$

A variant of this primal, without the connections to model calibration, has been studied by many authors in the context of optimal portfolio choice under exponential utility (see for instance [34, 46]).

The dual formulation reads

$$\begin{aligned} \widetilde{\text{MK}}_n^{\lambda*} \equiv & \inf_{(u_i(\cdot))_{1 \leq i \leq n}, (h_i(\cdot))_{1 \leq i \leq n}, \zeta \in \mathbb{R}_+} \sum_{i=1}^n \mathbb{E}^{\mathbb{P}^i} [u_i(S_i)] \\ & + \zeta \left(\lambda + \ln \mathbb{E}^{\mathbb{P}^0} [e^{\zeta^{-1} (c - \sum_{i=1}^n u_i(S_i) - \sum_{i=1}^n h_i(S_1, \dots, S_{i-1})(S_i - S_{i-1}))}] \right) \end{aligned} \quad (1.28)$$

Here the Lagrange multiplier $\zeta \in \mathbb{R}_+$ corresponds to the additional constraint $H(\mathbb{P}, \mathbb{P}^0) \leq \lambda$.

Theorem 1.4.3 (see A2) *Assume that the set $\{\mathbb{P} \in \mathcal{M}(\mathbb{P}^1, \dots, \mathbb{P}^n), H(\mathbb{P}, \mathbb{P}^0) \leq \lambda\}$ is non-empty. Let $c : \mathbb{R}_+^n \rightarrow [0, \infty)$ be an upper semi-continuous function so that*

$$c(s_1, \dots, s_n) \leq K \cdot (1 + |s_1| + \dots + |s_n|) \quad (1.29)$$

on \mathbb{R}_+^n for some constant K . Then there is no duality gap $\widetilde{\text{MK}}_n^{\lambda*} = \widetilde{\text{MK}}_n^\lambda$. The supremum is attained by the optimal measure \mathbb{P}^* given by the Gibbs density

$$\frac{d\mathbb{P}^*}{d\mathbb{P}^0} = \frac{e^{(\zeta^*)^{-1} (c - \sum_{i=1}^n u_i^*(S_i) - \sum_{i=1}^n h_i^*(S_i - S_{i-1}))}}{\mathbb{E}^{\mathbb{P}^0} [e^{(\zeta^*)^{-1} (c - \sum_{i=1}^n u_i^*(S_i) - \sum_{i=1}^n h_i^*(S_i - S_{i-1}))}]} \quad (1.30)$$

where $((u_i^*(\cdot))_{1 \leq i \leq n}, (h_i^*(\cdot))_{1 \leq i \leq n}, \zeta^*)$ achieves the infimum in (1.28).

The sagacious reader will note that a similar duality result is used for proving Talagrand's inequalities in optimal transport. Also the relative entropy can be replaced by any convex functional $f(\mathbb{P})$.

1.4.3 American options ♣

The martingale optimal transport problem in the case of a two-period Bermudan option with an exercise payoff $g(S)$ at t_1 and t_2 is defined as

$$\text{MK}_{\text{Ame}} \equiv \inf_{u_1, u_2, h_1, h_2} \mathbb{E}^{\mathbb{P}^1} [u_1(S_1)] + \mathbb{E}^{\mathbb{P}^2} [u_2(S_2)]$$

such that

$$\begin{aligned} u_1(s_1) + u_2(s_2) + h_1(s_1)(s_2 - s_1) &\geq g(s_1), \quad \forall s_1, s_2 \in \mathbb{R}_+ : \text{exercise at } t_1 \\ u_1(s_1) + u_2(s_2) + h_2(s_1)(s_2 - s_1) &\geq g(s_2), \quad \forall s_1, s_2 \in \mathbb{R}_+ : \text{exercise at } t_2 \end{aligned}$$

Here we have two delta functions $h_1(s_1)$ and $h_2(s_1)$ because our delta hedging strategy depends at t_1 on the value S_{t_1} and also if the payoff g has been exercised or not. The dual is then

$$\text{MK}_{\text{Ame}} = \sup_{p_1(s_1, s_2) \geq 0, p_2(s_1, s_2) \geq 0} \int g(s_1) p_1(ds_1, ds_2) + \int g(s_2) p_2(ds_1, ds_2)$$

such that

$$\begin{aligned} \int (s_2 - s_1)p_1(s_1, s_2)ds_2 = 0, \quad \int (s_2 - s_1)p_2(s_1, s_2)ds_2 = 0 \\ \int (p_1(s_1, s_2) + p_2(s_1, s_2))ds_2 = \mathbb{P}^1(s_1), \quad \int (p_1(s_1, s_2) + p_2(s_1, s_2))ds_1 = \mathbb{P}^2(s_2) \end{aligned}$$

In the particular case g convex, the optimal dual is

$$u_1^*(s_1) = 0, u_2^*(s_2) = g(s_2), h_1^*(s_1) = -g'(s_1), h_2^*(s_1) = 0$$

The optimal cost is reached by taking $p_1(ds_1, ds_2) = 0$ and p_2 an arbitrary martingale measure with marginals \mathbb{P}^1 and \mathbb{P}^2 . Indeed, we have $\text{MK}_{\text{Ame}} \leq \mathbb{E}^{\mathbb{P}^2}[g(S_{t_2})]$ and we deduce by weak duality the optimality of our guess. An open problem that we wish to address in our future research is the following: Apart for this trivial case, could this MOT be solved explicitly?

1.5 Link with Skorohod embedding problem [A3, S2]

1.5.1 Probabilistic framework

Let $\Omega \equiv \{\omega \in C([0, T], \mathbb{R}_+) : \omega_0 = 0\}$ be the canonical space equipped with the uniform norm $\|\omega\|_\infty \equiv \sup_{0 \leq t \leq T} |\omega_t|$, B the canonical process, \mathbb{P}_0 the Wiener measure, $\mathcal{F} \equiv \{\mathcal{F}_t\}_{0 \leq t \leq T}$ the filtration generated by B . S_0 is some given initial value in \mathbb{R}_+ , and we denote

$$S_t \equiv S_0 + B_t \quad \text{for } t \in [0, T].$$

For any \mathcal{F} -progressively measurable process σ and satisfying $\int_0^T \sigma_s^2 ds < \infty$, \mathbb{P}_0 -a.s., we define the probability measure on (Ω, \mathcal{F}) :

$$\mathbb{P}^\sigma \equiv \mathbb{P}_0 \circ (S^\sigma)^{-1} \quad \text{where } S_t^\sigma \equiv S_0 + \int_0^t \sigma_r dB_r, \quad t \in [0, T], \quad \mathbb{P}_0 - \text{a.s.}$$

Then S is a \mathbb{P}^σ -local martingale. We denote by \mathcal{P}_S the collection of all such probability measures on (Ω, \mathcal{F}) . The quadratic variation process $\langle S \rangle = \langle B \rangle$ is universally defined under any $\mathbb{P} \in \mathcal{P}_S$, and takes values in the set of all nondecreasing continuous functions. We shall then focus on the subset \mathcal{P}_∞ of \mathcal{P}_S consisting of all measures \mathbb{P} such that S is a \mathbb{P} -uniformly integrable martingale.

1.5.2 Model-free super-hedging problem

For all $\mathbb{P} \in \mathcal{P}_\infty$, we denote:

$$\mathbb{H}_{\text{loc}}^2(\mathbb{P}) \equiv \left\{ H \in \mathbb{H}^0(\mathbb{P}) : \int_0^T H_t^2 d\langle B \rangle_t < \infty, \quad \mathbb{P} - \text{a.s.} \right\},$$

Under the self-financing condition, for any portfolio process H , the portfolio value process

$$Y_t^H \equiv Y_0 + \int_0^t H_s \cdot dB_s, \quad t \in [0, T], \quad (1.31)$$

is well-defined \mathbb{P} -a.s. for every $\mathbb{P} \in \mathcal{P}_\infty$, whenever $H \in \mathbb{H}_{loc}^2$. Let ξ be an \mathcal{F}_T -measurable random variable. We introduce the subset of martingale measures:

$$\mathcal{P}_\infty(\xi) \equiv \{\mathbb{P} \in \mathcal{P}_\infty : \mathbb{E}^\mathbb{P}[\xi^-] < \infty\}.$$

The reason for restricting to this class of models is that, under the condition that $\mathbb{E}^\mathbb{P}[\xi^+] < \infty$, the hedging cost of ξ under \mathbb{P} is expected to be $-\infty$ whenever $\mathbb{E}^\mathbb{P}[\xi^-] = \infty$. As usual, in order to avoid doubling strategies, we introduce the set of admissible portfolios:

$$\mathcal{H}(\xi) \equiv \{H : H \in \mathbb{H}_{loc}^2 \text{ and } Y^H \text{ is a } \mathbb{P} - \text{supermartingale for all } \mathbb{P} \in \mathcal{P}_\infty(\xi)\}.$$

In addition to the continuous-time trading, we assume that the investor can take static positions in T -Vanilla options. The T -European option defined by the payoff $\lambda(S_T)$ has an un-ambiguous market price defined by

$$\mu(\lambda) = \int \lambda d\mu$$

We denote below $\xi^\lambda \equiv \xi - \lambda$. The set of European payoffs which may be used by the hedger are naturally taken in the set

$$\Lambda^\mu(\xi) \equiv \left\{ \lambda \in \Lambda^\mu : \sup_{\mathbb{P} \in \mathcal{P}} \mathbb{E}^\mathbb{P}[(\xi^\lambda)^+] < \infty \right\}, \text{ where } \Lambda^\mu \equiv \left\{ \lambda : \lambda \in L^1(\mu) \right\}. \quad (1.32)$$

The superreplication upper bound is defined by:

$$U^\mu(\xi) \equiv \inf \left\{ Y_0 : \exists \lambda \in \Lambda^\mu(\xi) \text{ and } H \in \mathcal{H}(\xi^\lambda), \bar{Y}_T^{H,\lambda} \geq \xi, \mathbb{P} - \text{a.s. for all } \mathbb{P} \in \mathcal{P}_\infty(\xi^\lambda) \right\}, \quad (1.33)$$

where $\bar{Y}^{H,\lambda}$ denotes the portfolio value of a self-financing strategy with continuous trading H in the primitive securities, and static trading λ in the T -Vanillas:

$$\bar{Y}_T^{H,\lambda} \equiv Y_0 + \int_0^T H_s dB_s - \mu(\lambda) + \lambda(S_T), \quad (1.34)$$

indicating that the investor has the possibility of buying at time 0 any derivative security with payoff $\lambda(S_T)$ for the price $\mu(\lambda)$. $U^\mu(\xi)$ is an upper bound on the price of ξ necessary for absence of strong (model-independent) arbitrage opportunities: selling ξ at a higher price, the hedger could set up a portfolio with a negative initial cost and a non-negative payoff under any market scenario. The next result gives a dual formulation of the robust superhedging:

Proposition 1.5.1 (see A3) *Assume that $\sup_{\mathbb{P} \in \mathcal{P}_\infty} \mathbb{E}^\mathbb{P}[\xi^+] < \infty$. Then:*

$$U^\mu(\xi) = \inf_{\lambda \in \Lambda^\mu(\xi)} \sup_{\mathbb{P} \in \mathcal{P}_\infty} \left\{ \mu(\lambda) + \mathbb{E}^\mathbb{P}[\xi - \lambda(S_T)] \right\}.$$

It could be tempting to permute the inf and the sup - we obtain

$$U^\mu(\xi) \stackrel{?}{=} \sup_{\mathbb{P} \in \mathcal{P}_\infty} \inf_{\lambda \in \Lambda^\mu(\xi)} \left\{ \mu(\lambda) + \mathbb{E}^\mathbb{P}[\xi - \lambda(S_T)] \right\} = \sup_{\mathbb{P} \in \mathcal{P}_\infty, S_T \sim \mu} \mathbb{E}^\mathbb{P}[\xi]$$

So providing we could justify this minimax argument, our robust superhedging is connected to a martingale optimal transport: we maximize the cost $\mathbb{E}^\mathbb{P}(\xi)$ over the space of martingale measure with T -marginal μ and $\mathbb{P}_0 = \delta_{S_0}$. If we assume that the payoff $\xi(S_T, M_T, m_T, \langle S \rangle_T)$ depends only on the spot S_T , the running maximum M_T , minimum m_T or the quadratic variation $\langle S \rangle_T$ at T , by doing a stochastic time change, U^μ can be framed as a constrained perpetual American options:

$$U^\mu(\xi) = \sup_{\tau \in \mathcal{T} : B_\tau \sim \mu} \mathbb{E}^\mathbb{P} \left[\xi \left(B_\tau, \max_{0 \leq s \leq \tau} B_s, \min_{0 \leq s \leq \tau} B_s, \tau \right) \right]$$

This problem corresponds then to the determination of an (optimal) stopping time τ^* such that $B_{\tau^*} \sim \mu$. This is a Skorohod embedding problem (SEP). The use of SEP for deriving model-independent bounds for exotic options consistent with T -vanillas was pioneered by D. Hobson in the case of lookback options [65] and then by B. Dupire in the case of options on variance [44].

Note that this approach should give a remarkable solution of SEP as the associated martingale S_t has the property to maximize the expectation of the payoff ξ over the class of arbitrage-free models calibrated to T -Vanilla options. This link between optimal transport and SEP has been recently explored in [14].

Although most known solutions of SEP share this optimality property, this property was checked in a second step in the original papers on SEP. Our approach consisting in solving this problem of perpetual American options implies at the beginning that our solution to SEP satisfies an optimality property. In the next section, we show on the simple example of a payoff $\xi = g(M_T)$ depending on an increasing function on the running maximum M_T how to reproduce Azema-Yor's solution [11]. Although the derivation of this solution is simple using martingale techniques (mainly a stopped Azema-Yor's martingale), we highlight that our method is straightforward and can easily handle more complicated examples. The determination of a SEP synthetists in the analytical resolution of a one (eventually two)-dimensional time-homogeneous obstacle problem. In particular, following the main steps sketched below, one can show that the optimality of Azema-Yor's solution is still valid with a payoff $\xi = g(S_T, M_T)$ where $s \rightarrow \frac{\partial_m g(s, m)}{m-s}$ is a nondecreasing function for all $s < m$ (see [S2]). Our approach can also tackle the multi-marginals SEP (see [S2] and section below).

1.5.3 Azema-Yor's solution revisited [A3]

The aim of this section is to reproduce within our framework Azema-Yor's solution to the Skorohod embedding problem which gives the optimal bound for options written on the maximum of an underlying: $\xi = g(M_T)$ with $M_T = \max_{s \leq T} S_s$. g is a nondecreasing function. Here we do not reproduce our proof in [A3] but instead sketch the derivation highlighting the main ideas and the novelty of our approach. We note

$$u^\lambda(s, m) \equiv \sup_{\tau \in \mathcal{T}} \mathbb{E}[g^\lambda(B_\tau^*) | B_0 = s, B_0^* = m]$$

with $B_\tau^* \equiv \max_{0 \leq s \leq \tau} B_s$ and $g^\lambda(s, m) \equiv g(m) - \lambda(s)$. Then, $U^\mu(\xi)$ can be written as

$$U^\mu(\xi) = \inf_{\lambda \in \Lambda^\mu(\xi)} \{ \mu(\lambda) + u^\lambda(S_0, S_0) \}. \quad (1.35)$$

u^λ satisfies in the viscosity sense the variational equation

$$\max(g^\lambda(s, m) - u^\lambda(s, m), \partial_s^2 u^\lambda(s, m)) = 0, \quad s \leq m \quad (1.36)$$

with the Neumann condition $\partial_m u^\lambda(m, m) = 0$. As an ansatz for the optimal stopping time, we take

$$\tau_* = \inf\{t > 0 : B_t \leq \Psi_\lambda(B_t^*)\}$$

with $\Psi_\lambda(\cdot) : [S_0, \infty] \rightarrow \mathbb{R}$. u^λ satisfies the following system:

$$\begin{aligned} \partial_s^2 u^\lambda(s, m) &= 0, \quad \Psi_\lambda(m) < s < m \\ \partial_m u^\lambda(m, m) &= 0 : \text{(normal reflection)} \\ u^\lambda(\Psi_\lambda(m), m) &= g(m) : \text{(instantaneous stopping)} \end{aligned}$$

that we complete by the additional requirement (the delta is smooth at the boundary)

$$\partial_s u^\lambda(\Psi_\lambda(m), m) = 0 : \text{(smooth fit)}$$

This can be solved explicitly:

$$u^\lambda(s, m) = g(m) + \int_{\Psi_\lambda(m)}^{s \vee \Psi_\lambda(m)} (s - x) \lambda''(x) dx \quad (1.37)$$

In order to satisfy the obstacle problem, we should impose that λ is a convex function. Due to the normal reflection condition, the optimal stopping boundary $\Psi_\lambda(m)$ is the solution of the ODE:

$$\Psi'_\lambda(m) \lambda''(\Psi_\lambda(m)) (m - \Psi_\lambda(m)) = g'(m) \quad (1.38)$$

which should stay below and never hit the diagonal in \mathbb{R} , i.e. $\Psi_\lambda(m) < m$ for all m . Note that as no initial conditions are specified, the solution of this ODE (and the solution of the variational equation (1.36)) is not unique. Finally, by restricting the infimum over λ to the smaller set of λ convex, we get that

$$U^\mu(\xi) \leq \inf_{\lambda(\cdot) \text{ conv}} \left(g(S_0) + \int_{\Psi_\lambda(S_0)}^{S_0} (S_0 - x) \lambda''(x) dx + \int_0^\infty \lambda''(x) C^\mu(T, x) dx \right)$$

with $C^\mu(T, x) \equiv \mathbb{E}^\mu[(S_T - x)^+]$. By doing the change of variable $x = \Psi_\lambda(y)$ and plugging the expression (1.38) for $\lambda''(\cdot)$ in the above equation, we get

$$\inf_{\Psi_\lambda(\cdot)} g(S_0) + \int_0^\infty \{(S_0 - \Psi_\lambda(y)) \mathbf{1}_{\Psi_\lambda(y) \in [\Psi_\lambda(S_0), S_0]} + C^\mu(T, \Psi_\lambda(y))\} \frac{g'(y)}{y - \Psi_\lambda(y)} dy$$

Pointwise minimization gives

$$(S_0 - y) \mathbf{1}_{\Psi_\lambda(y) \in [\Psi_\lambda(S_0), S_0]} + (y - \Psi_\lambda(y)) \partial_2 C^\mu(T, \Psi_\lambda(y)) + C^\mu(T, \Psi_\lambda(y)) = 0$$

By setting $y = \Psi_\lambda^{-1}(x)$, we get

$$(S_0 - \Psi_\lambda^{-1}(x))1_{x \in [\Psi_\lambda(S_0), S_0]} + (\Psi_\lambda^{-1}(x) - x)\partial_2 C^\mu(T, x) + C^\mu(T, x) = 0$$

for which the solution is

$$\Psi_\mu^{-1}(x) = \frac{x\partial_2 C^\mu(T, x) - C^\mu(T, x)}{\partial_2 C^\mu(T, x)} = \frac{\mathbb{E}^\mu[S_T 1_{S_T \geq x}]}{\mathbb{E}^\mu[1_{S_T \geq x}]} \quad (1.39)$$

This reproduces Azéma-Yor's solution [11]. As a byproduct, we obtain a path-wise inequality corresponding to the interpretation of the upper bound in terms of robust super-replication:

$$U^\mu(g) + \int_0^T H_t dS_t + \lambda^*(S_T) - \mu(\lambda^*) \geq g(M_T), \quad \mathbb{P} - \text{a.s. for all } \mathbb{P} \in \mathcal{P}_\infty(\xi^{\lambda^*})$$

with

$$H_s = \partial_s u^{\lambda^*}(S_t, M_t), \quad \lambda^{*''}(\Psi_\mu(m))\Psi_\mu''(m) = \frac{g'(m)}{m - \Psi_\mu(m)}$$

This could be justified from a classical verification argument using Itô's formula.

A similar approach can be used for payoffs depending on the spot, minimum and maximum (in preparation). As seen previously, Azema-Yor's solution is still optimal if we include the information of all Vanillas with maturities less than T in the case of barycenter functions $\Psi_{\text{AY}, \mathbb{P}^t}(x)$ increasing in time. The optimal upper bound U^μ depends then only on the (terminal) T -Vanillas. Below, we explain how to extend this result when this assumption is not satisfied.

1.5.4 Multi-marginals [S2]

The key ingredient for the solution of the present n -marginals Skorohod embedding problem turns out to be the following minimization problem:

$$C(m) \equiv \min_{\zeta_1 \leq \dots \leq \zeta_n \leq m} \sum_{i=1}^n \left(\frac{C(t_i, \zeta_i)}{m - \zeta_i} - \frac{C(t_i, \zeta_{i+1})}{m - \zeta_{i+1}} 1_{\{i < n\}} \right) \quad \text{for all } m \geq S_0, \quad (1.40)$$

where $C(t_i, x) = \mathbb{E}^{\mathbb{P}^i}[(S_{t_i} - x)^+]$. We report the following result:

Theorem 1.5.2 (see S2) *Let g be a non-decreasing function and assume that $\mathbb{P}^1 \leq \dots \leq \mathbb{P}^n$. Let $\zeta_1^*(m), \dots, \zeta_n^*(m)$ be a solution to (1.40) for a fixed m . Then,*

$$U^\mu(\xi) \leq U \equiv g(S_0) + \sum_{i=1}^n \int_{S_0}^{\infty} \left(\frac{c_i(\zeta_i^*(m))}{m - \zeta_i^*(m)} - \frac{c_i(\zeta_{i+1}^*(m))}{m - \zeta_{i+1}^*(m)} 1_{\{i < n\}} \right) dg(m). \quad (1.41)$$

Moreover, there exist $\lambda \in \Lambda_n^\mu$, and trading strategies $H = H^{\text{stock}} + H^{\text{fwd}} \in \mathcal{H}(\xi^\lambda)$, explicitly given in [S2], such that $U = g(S_0) + \sum_{i=1}^n \mathbb{E}^{\mathbb{P}^i}[\lambda_i(S_{t_i})]$ and

$$U + \sum_{i=1}^n \lambda_i(S_{t_i}) - \sum_{i=1}^n \mathbb{E}^{\mathbb{P}^i}[\lambda_i(S_{t_i})] + \sum_{i=1}^n H_{t_{i-1}}(S_{t_i} - S_{t_{i-1}}) \geq g(M_T) \quad \text{for all } \omega \in \Omega. \quad (1.42)$$

Assume further that $\mathbb{P}^1, \dots, \mathbb{P}^n$ satisfy Assumption A in [94]. Then, equality holds in (1.41).

1.6 Stochastic inequalities

Mass transportation provides a power tool to study some functional inequalities with geometric content. Let us mention the optimal Sobolev inequality, entropy-entropy production inequalities, transportation inequalities. In this section, we explain how martingale mass transportations can provide similar results. We illustrate this on Doob's inequality and Burkholder-Davis-Gundy's inequality for $p = 1$. As a byproduct, we obtain path-wise inequalities. Alternative approaches are presented in [2, 13, 14].

1.6.1 Doob's inequality revisited

Let S_t be a càdlàg positive martingale starting at S_0 at $t = 0$. Doob's inequality states that

$$\mathbb{E}[M_T^p] \leq S_0^p + \left(\frac{p}{p-1}\right)^{\frac{1}{p}} \mathbb{E}[S_T^p], \quad \forall p > 1$$

The derivation of this inequality (with the optimal constant $\left(\frac{p}{p-1}\right)$) can be framed into the computation of a robust upper bound of the payoff M_T^p within the class of arbitrage-free models calibrated to T -power options with payoff S_T^p (this section originates for an internal presentation at Société Générale). Following our discussion on Azema-Yor's solution, we define

$$U^\mu = \inf_{\lambda \in \mathbb{R}} u^\lambda(S_0, S_0) + \lambda \mathbb{E}[S_T^p]$$

where $u^\lambda(s, m) = \sup_{\tau \in \mathcal{T}} \mathbb{E}_{s,m}[(B_\tau^*)^p - \lambda B_\tau^p | B_0 = s, B_0^* = m]$. u^λ is a viscosity solution of the variational inequality:

$$\max(m^p - \lambda s^p - u^\lambda(s, m), \partial_s^2 u^\lambda(s, m)) = 0, \quad s \leq m$$

with the Neumann condition $\partial_m u^\lambda(m, m) = 0$. Using a similar ansatz for the optimal stopping time as in Azema-Yor's solution, we derive a solution (1.37) and Ψ satisfies (1.38) with $\lambda(x) = \lambda x^p$. Solutions are $\psi^\lambda(m) = c m$ with $\lambda(p-1)c^{p-1}(1-c) = 1$ and $c < 1$. This implies that

$$U^\mu \leq S_0^p + \frac{1}{(p-1)c^{p-1}(1-c)} \mathbb{E}[S_T^p]$$

Taking the infimum over $c \in (0, 1)$, we obtain

$$U^\mu \leq S_0^p + \left(\frac{p}{p-1}\right)^p \mathbb{E}[S_T^p]$$

The supremum was reached for $c^* = \frac{(p-1)}{p}$ and $\lambda^*(p-1)(c^*)^{p-1}(1-c^*) = 1$. Finally we check that our guess stopping time $\tau^* = \inf\{t > 0 : B_t \leq c^* B_t^*\}$ is optimal. From our stochastic control approach, we obtain the path-wise inequality:

$$U^\mu + \int_0^T \partial_s u^{\lambda^*}(S_s, M_s) dS_s + \lambda^*(S_T^p - \mathbb{E}[S_T^p]) \geq M_T^p, \quad \mathbb{P} - \text{a.s. for all } \mathbb{P} \in \mathcal{P}_\infty$$

A similar inequality was provided recently using an alternative approach [2].

1.6.2 Burkholder-Davis-Gundy inequality revisited ♣

Another classical result in stochastic analysis is the Burkholder-Davis-Gundy (in short BDG) inequality which states that for any $p \geq 0$ there exists universal constants $c_1(p)$ and $c_2(p)$ such that for all local martingales S with $S_0 \equiv 0$ and stopping time T , the following inequality holds

$$c_1(p)\mathbb{E}[V_T^{\frac{p}{2}}] \leq \mathbb{E}[M_T^p] \leq c_2(p)\mathbb{E}[V_T^{\frac{p}{2}}]$$

Here $V_t \equiv \langle S \rangle_t$ denotes the quadratic variation and $M_t \equiv \max_{0 \leq s \leq t} S_s$ the running maximum. For continuous local martingales, this statement holds for all $p > 0$.

As stated in Peskir, Shiryaev [101], “the question of finding the best possible values for $c_1(p)$ and $c_2(p)$ appears to be of interest. Its emphasis is not so much on having these values but more on finding a method of proof which can deliver them.” In the case where $\mathbb{E}[M_T^p]$ is replaced by $\mathbb{E}[S_T^p]$, the optimal constants can be computed by the method of time change as described in [101], Section 10.

For the BDG inequalities, only partial results are known: For $p = 2$, the best constants are $c_1(2) = 1$ and $c_2(2) = 4$ (See [101] Section 20). One is deduced from Doob’s inequality and the other from the stopping time $\tau = \inf\{t > 0 : |B_t| = 1\}$ with B a one-dimensional Brownian motion. Quite recently, for $p = 1$, Burkholder [27] obtains $c_1(1) = 1/\sqrt{3}$ and Osekowski [96] $c_2(1) \simeq 1.4658$. Pecatti and Yor [100] shows that if we consider the subclass of continuous martingale, the constant $1/\sqrt{3}$ is no longer optimal and could be replaced by $1/\sqrt{2}$. In this section, we reproduce the optimal constant for $p = 1$ (for a refined BDG version) and provide also an (optimal) path-wise version of this inequality. The case $p > 1$ is still in progress.

Theorem 1.6.1 (Optimal BDG $p = 1$, see also [96] Theorem 4.1) *For all stopping times T , for all local continuous martingales S*

$$\mathbb{E}[M_T] \leq S_0 + c_2(1)\mathbb{E}[\sqrt{V_T}] \tag{1.43}$$

with $c_2(1) \simeq 1.30693$ is the optimal constant, solution of

$$e^{\frac{t_*^2}{2}} = t_* \int_0^{t_*} e^{\frac{u^2}{2}} du$$

This bound corresponds to the path-wise inequality

$$\begin{aligned} S_0 + c_2(1)\mathbb{E}[\sqrt{V_T}] + c_2(1) \left(\sqrt{V_T} - \mathbb{E}[\sqrt{V_T}] \right) \\ + \int_0^T \left(-\frac{M' \left(\frac{S_s - M_s}{\sqrt{V_s}} \right)}{M'(-c_2(1))} + 1 \right) dS_s \geq M_T, \quad \mathbb{P} - a.s. \text{ for all } \mathbb{P} \in \mathcal{P}_\infty \end{aligned} \tag{1.44}$$

with $M(x) \equiv M(-\frac{1}{2}, \frac{1}{2}, \frac{x^2}{2})$. $M(a, b, c)$ is Kummer’s confluent hypergeometric function.

Note that if we take the expectation of the inequality (1.44), we obtain (1.43). The above path-wise inequality (1.44) has a nice interpretation in mathematical finance. The BDG inequality

(for $p = 1$) corresponds to the computation of a model-independent bound for a lookback payoff M_T with maturity T consistent with the (market) price of (normal) volatility swap with payoff $\sqrt{V_T}$. This bound can be super-replicated by a strategy involving at $t = 0$ a cash amount $S_0 + c_2(1)\mathbb{E}[\sqrt{V_T}]$, a static hedge of the payoff $\sqrt{V_T}$ with weight $c_2(1)$ and finally a delta hedging $h_s = A_1(1)M' \left(\frac{S_s - M_s}{\sqrt{V_s}} \right) + \frac{1}{\sqrt{2}} \left(\frac{S_s}{\sqrt{V_s}} \right)$.

Our derivation follows closely our section on Azema-Yor's solution and Doob's inequality. Here we consider the general case $p \geq 1$ and specialized to $p = 1$ later one. We have that for all stopping time T ,

$$\mathbb{E}[M_T^p] = \mathbb{E}[M_T^p - \lambda V_T^{\frac{p}{2}}] + \lambda \mathbb{E}[V_T^{\frac{p}{2}}] \leq \inf_{\lambda \in \mathbb{R}} \sup_{\tau} \mathbb{E}[M_\tau^p - \lambda V_\tau^{\frac{p}{2}}] + \lambda \mathbb{E}[V_T^{\frac{p}{2}}]$$

We set below $u^\lambda(s, m, v) \equiv \sup_{\tau} \mathbb{E}_{s, m, v}[M_\tau^p - \lambda V_\tau^{\frac{p}{2}}]$ and

$$\mathbb{E}[M_T^p] \leq \inf_{\lambda \in \mathbb{R}} u^\lambda(S_0, S_0, 0) + \lambda \mathbb{E}[V_T^{\frac{p}{2}}]$$

$u^\lambda(s, m, v)$ solves the dynamic programming equation

$$\max \left(m^p - \lambda v^{\frac{p}{2}} - u^\lambda, \frac{1}{2} u_{ss}^\lambda + u_v^\lambda \right) = 0, \quad s \leq m$$

subject to the Neumann condition $u_m(m, m, v) = 0$. As stated in Peskir and Shiryaev [101], section 20, this non-linear problem is inherently three-dimensional and quite difficult to solve. In fact this PDE can be converted into a two-dimensional PDE, quite similar to the variational equation (1.36) appearing in Azema-Yor's solution. We simplify this equation with the change of variables $u^\lambda(s, m, v) = v^{\frac{p}{2}} U^\lambda \left(x \equiv \frac{s}{\sqrt{v}}, y \equiv \frac{m}{\sqrt{v}}, z \equiv \frac{1}{\sqrt{v}} \right)$. It is clear that the function $u^\lambda(s, m, v)$ should scale as α^p when $x \rightarrow \alpha x$, $m \rightarrow \alpha m$ and $v \rightarrow \alpha^2 v$. This scaling symmetry implies that the function U^λ is independent of z and therefore solves a *two-dimensional* PDE:

$$\max \left(y^p - \lambda - U^\lambda(x, y), U_{xx}^\lambda + pU^\lambda - xU_x^\lambda - yU_y^\lambda \right) = 0, \quad x \leq y, \quad y \in \mathbb{R}_+$$

subject to the Neumann condition $U_y^\lambda(y, y) = 0$. The derivation of the optimal BDG inequality is then achieved by guessing a solution of the above equation of the form

$$\begin{aligned} U^\lambda(x, y) &= y^p - \lambda, \quad \forall x \in (-\infty, \psi_\lambda(y)] \\ U_{xx}^\lambda + pU^\lambda - xU_x^\lambda - yU_y^\lambda &= 0, \quad \forall x \in [\psi_\lambda(y), y] \\ U_y^\lambda(y, y) &= 0 \\ U^\lambda(x, y) &\geq y^p - \lambda, \quad \forall x \in [\psi_\lambda(y), y] \\ U_x^\lambda(\psi_\lambda(y), y) &= 0 \end{aligned}$$

where the optimal stopping time is

$$\tau_{\psi_\lambda} = \inf \left\{ t > 0 : S_t \leq \sqrt{V_t} \psi_\lambda \left(\frac{M_t}{\sqrt{V_t}} \right) \right\}$$

The solution for $p = 1$ is

$$U^\lambda(x, y) = -\frac{M(x - y)}{M'(-c_2(1))} + x$$

and $\lambda^* = c_2(1)$ is solution of $M(-c_2(1)) = 0$. The super-replication property could be justified from a classical verification argument using Itô's formula. The analysis of the case $p > 1$ deserves further study.

Chapter 2

Asymptotics of implied volatility: A geometric approach

In this chapter, we consider short-time and large-strike asymptotics of implied volatility for local and time-homogeneous stochastic volatility model. Our approach, started in 2004, uses powerful geometrical methods such as short-time heat kernel expansion on Riemannian manifolds. As a striking application, this approach allows to derive a short-time arbitrage-free smile for the well-known SABR model making a beautiful connection with the geometry of the Poincaré plane. Our approach has generated an intensive research of both a theoretical and practical nature. Let us mention two international conferences on this subject <http://www.math.nyu.edu/~laurence/vienna-Sabr-bis.htm> and <http://www2.imperial.ac.uk/~ajacquie/WorkshopLDP/index.html>.

2.1 Motivation: risk model issues

As a full-blown model generates by construction arbitrage-free Vanillas, it is quite natural to (a) choose a basic model, (b) parameterize a Vanilla smile using the model's parameters. The model itself is only used as a fitting tool. In 2002 Hagan-al [61] derived an approximate expression for the smile of the SABR model that holds for short maturities and strikes near the money. The SABR formula has then become a standard in fixed income markets. The SABR model, depending on 4 parameters α, β, ρ, ν , is a stochastic volatility model (in short SVM) specified by the SDE:

$$\begin{aligned}dS_t &= C(S_t)a_t dW_t, & C(x) &= x^\beta \\ da_t &= \nu a_t dZ_t, & a_0 &= \alpha, & d\langle W, Z \rangle_t &= \rho dt\end{aligned}$$

The approximate implied volatility for a strike K and a maturity T in the short-time limit at zero order in the maturity (i.e. independent of T) and for strikes near the money is (see last equation

in [61])

$$\sigma_{\text{BS}}(T, K) = \frac{\alpha \ln \frac{S_0}{K}}{\int_K^{S_0} \frac{du}{C(u)}} \frac{z}{x(z)} + O(T)$$

where $z = \frac{\nu}{\sigma} \frac{S_0 - S}{C(\sqrt{SS_0})}$ and $x(z) = \ln \left(\frac{\sqrt{1 - 2\rho z + z^2} + z - \rho}{1 - \rho} \right)$. In 2004, just beginning in the model validation group at Barclays Capital, I was asked to audit this formula. It turned out that once calibrated on market Vanillas, the T -marginal implied by the SABR formula can become negative (highlighting the presence of arbitrage) for large maturity $T \geq 10$ years (standard value in fixed income markets) and for small and large strikes. In particular for large strikes, we obtain for the (first order) SABR formula ($\beta \in [0, 1)$)

$$\sigma_{\text{BS}}(T, K) \sim_{K \rightarrow \infty} e^{\beta \ln K}$$

This asymptotic behavior, applied to the case $\beta = 1$, is in contradiction with a result by Lee [86] that states that the large-strikes asymptotics cannot grow faster than linearly in $\ln K$ with a slope depending on the existence of higher moments. Since the high-strike smile is arbitrageable, the price of Constant Maturity Swap obtained by static replication from swaption smiles is infinite. From these risk model issues, I have started (a) correcting the short-time asymptotics SABR formula (specially the small and large-strike behavior) and (b) deriving a similar formula for arbitrary (time-homogeneous) SVM. For example, one would like to include a mean-reversion drift in the volatility (called λ -SABR in [A11]). Hagan-al derivation is not straightforward and require performing a multiple of (judicious) change of variables. A generalization to other SVMs seemed quite difficult at that time. It turns out that the right framework is short-time heat kernel expansions on Riemannian manifold and the (corrected version of) SABR formula is a reminiscence of the geometry of geodesic curves in the Poincaré hyperbolic plane (see Figure 2.1). A similar approach (albeit less geometrical and only restricted to the zero-order) was performed by Avellaneda and al [10] for a basket smile in the case of a multi-local volatility model and by Berestycki and al [18] independently at the same time. This approach is extensively covered in my book [B1]. Then, this topics has generated a lot of research. One can cite some extensions for time-dependent local and stochastic volatility models (see e.g. [56, 58]) and some other alternative methods (as the large deviation approach, Watanabe-Malliavin calculus) that can take care of some stochastic volatility models that can not be covered by our geometric approach (see e.g. [50]). For example, the Heston model is connected to an incomplete manifold (see Remark 2.2.2).

2.2 From risk model issues to geometry [B1,A10,A11,A12]

A (time-homogeneous) one-factor SVM depends on two SDEs, one for the asset S and one for the volatility a . Let us denote $X = (X^i)_{i=1,2} = (S, a)$, with initial conditions $\alpha = (\alpha^i)_{i=1,2} = (S_0, a_0)$. These processes X_t^i satisfy the following SDEs

$$dX_t^i = b^i(X_t)dt + \sigma^i(X_t)dW_t^i, \quad d\langle W^i, W^j \rangle_t = \rho_{ij}dt \quad (2.1)$$

with the initial condition $X_0^i = \alpha^i$ and $b^1 = 0$. Here $[\rho_{ij}]_{i,j=1,2}$ is the (constant) correlation matrix.

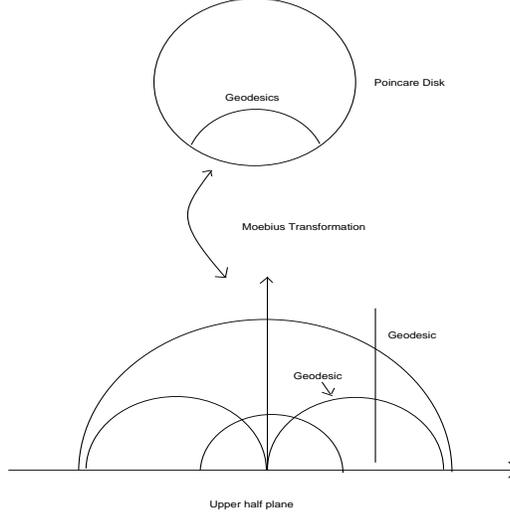


Figure 2.1: Poincaré disk \mathcal{D} and upper half-plane \mathbb{H}^2 with some geodesics. In the upper half-plane, the geodesics correspond to vertical lines and to semi-circles centered on the horizon $\Im(z) = 0$ and in \mathcal{D} the geodesics are circles orthogonal to \mathcal{D} .

From [42], we have that X_t^1 admits the same marginals as the one-dimensional SDE

$$d\bar{X}_t^1 = \Sigma(t, \bar{X}_t^1) dW_t$$

if the square of the local volatility function $\Sigma(\tau, S)$ is equal to the square of the stochastic volatility $d\langle \ln X_t^1 \rangle / dt$ when the forward X_τ^1 is fixed to S :

$$\begin{aligned} \Sigma^2(\tau, S) &= \mathbb{E}[(\sigma^1(S_\tau, a_\tau))^2 | S_\tau = S] \\ &= \frac{\int_0^\infty (\sigma^1(S, a))^2 p(S, a | S_0, \alpha) da}{\int_0^\infty p(S, a | S_0, \alpha) da} \end{aligned}$$

where $p(\tau, x | \alpha)$ - assumed to exist if the diffusion is hypo-elliptic - is the conditional probability. In order to obtain an asymptotic expression for the implied volatility, we will (a) derive a short-time asymptotic expansion for the local volatility, then (b) use a short-time asymptotic relation between the implied volatility and the local volatility $\sigma(\tau, S)$ (see [17] for the zero-order term and [B1] for the first order correction). The density p satisfies the backward Kolmogorov equation (the time has been reverted $\tau = T - t$)

$$\frac{\partial p}{\partial \tau} = b^i \partial_i p + g^{ij} \partial_{ij} p, \quad (i, j) = S, a \quad (2.2)$$

Here $g^{ij} \equiv \frac{1}{2} \rho_{ij} \sigma^i \sigma^j$ and $\partial_i \equiv \frac{\partial}{\partial \alpha^i}$. In (2.2), we have used the Einstein summation convention. We will adopt this Einstein convention throughout this chapter. In the next section, we will show

how to derive an asymptotic short-time conditional probability p for any multi-dimensional time-homogeneous stochastic volatility models (2.1), not necessarily two-dimensional, using the heat kernel expansion on a Riemannian manifold. As these tools are not commonly used (and known) by people working in mathematical finance and probability theory, we give below a short reminder of the key notions.

2.2.1 Heat kernel on a Riemannian manifold

In this section, PDE (2.2) will be interpreted as a heat kernel on a general smooth n -dimensional manifold M (here we have that $i, j = 1 \cdots n$) without boundary, endowed with the metric g_{ij} (see [B1], Chapter 4 for details). Note that the coordinates $\{\alpha^i\}$ (resp. $\{x^i\}$) will be noted $\{x^i\}$ (resp. $\{y^i\}$) below in order to be consistent with our (geometric) notation.

The inverse of the metric g^{ij} is defined by

$$g^{ij} = \frac{1}{2} \rho_{ij} \sigma^i \sigma^j$$

and the metric (ρ^{ij} inverse of ρ_{ij} , i.e., $\rho^{ij} \rho_{jk} = \delta_k^i$)

$$g_{ij} = 2 \frac{\rho^{ij}}{\sigma^i \sigma^j} \quad (2.3)$$

We denote $g \equiv \det[g_{ij}]$. The differential operator

$$D = b^i \partial_i + g^{ij} \partial_{ij}$$

which appears in (2.2) is a second-order elliptic operator of Laplace type. We can then show that there is a unique connection ∇ on \mathcal{L} , a line bundle over M , and Q a unique smooth section of $End(\mathcal{L}) = \mathcal{L} \otimes \mathcal{L}^*$ such that

$$\begin{aligned} D &= g^{ij} \nabla_i \nabla_j + Q \\ &= g^{-\frac{1}{2}} (\partial_i + \mathcal{A}_i) g^{\frac{1}{2}} g^{ij} (\partial_j + \mathcal{A}_j) + Q \end{aligned}$$

Using this connection ∇ , (2.2) can be written in the covariant way, i.e.,

$$\frac{\partial}{\partial \tau} p(\tau, x|y) = Dp(\tau, x|y) \quad (2.4)$$

If we take $\mathcal{A}_i = 0$ and $Q = 0$ then D becomes the Laplace-Beltrami operator $\Delta = g^{-\frac{1}{2}} \partial_i \left(g^{\frac{1}{2}} g^{ij} \partial_j \right)$. We may express the connection \mathcal{A}^i and Q as a function of the drift b^i and the metric g_{ij} by identifying in (2.4) the terms ∂_i and ∂_{ij} with those in (2.2). We find

$$\mathcal{A}^i = \frac{1}{2} \left(b^i - g^{-\frac{1}{2}} \partial_j \left(g^{1/2} g^{ij} \right) \right) \quad (2.5)$$

$$Q = g^{ij} (\mathcal{A}_i \mathcal{A}_j - b_j \mathcal{A}_i - \partial_j \mathcal{A}_i) \quad (2.6)$$

Note that the Latin indices $i, j \dots$ can be lowered or raised using the metric g_{ij} or its inverse g^{ij} . For example $\mathcal{A}_i = g_{ij}\mathcal{A}^j$ and $b_i = g_{ij}b^j$. The components $\mathcal{A}_i = g_{ij}\mathcal{A}^j$ define a local one-form $\mathcal{A} = \mathcal{A}_i dx^i$. We deduce that under a change of coordinates $x^{i'}(x^j)$, \mathcal{A}_i undergoes the co-vector transformation $\mathcal{A}_{i'}\partial_i x^{i'} = \mathcal{A}_i$. Note that the components b^i don't transform as a vector. This results from the fact that the SDE (2.1) has been derived using the Itô calculus and not the Stratonovich one.

Next, let's introduce the Christoffel's symbol Γ_{ij}^k which depends on the metric and its first derivatives

$$\Gamma_{ij}^k = \frac{1}{2}g^{kp}(\partial_j g_{ip} + \partial_i g_{jp} - \partial_p g_{ji}) \quad (2.7)$$

(2.5) can be re-written as

$$\mathcal{A}^i = \frac{1}{2}(b^i - g^{pq}\Gamma_{pq}^i)$$

Note that if we define

$$p' = e^{\chi(\tau,x) - \chi(\tau=0,x=\alpha)} p \quad (2.8)$$

then p' satisfies the same equation as p (2.4) but with

$$\mathcal{A}'_i \equiv \mathcal{A}_i - \partial_i \chi \quad (2.9)$$

$$Q' \equiv Q + \partial_\tau \chi \quad (2.10)$$

The transformation (2.8) is called a *gauge transformation*. The reader should be aware that the transformation (2.9) only applies to the connection \mathcal{A}_i with lower indices. The constant phase $e^{\chi(x=\alpha,\tau=0)}$ has been added in (2.8) so that p and p' satisfy the same boundary condition at $\tau = 0$. Mathematically, (2.8) means that p is a section of the line bundle \mathcal{L} and when we apply a (local) Abelian gauge transformation, this induces an action on the connection \mathcal{A} (2.9). In particular, if the one-form \mathcal{A} is exact, meaning that there exists a smooth function χ such that $\mathcal{A} = d\chi$ then the new connection \mathcal{A}' in (2.9) vanishes. See chapter 4 in [B1] for an interpretation of this transformation as a Girsanov change of measure.

The asymptotic resolution of the heat kernel (2.4) in the short time limit is an important problem in Theoretical Physics and in Mathematics. In Physics, it corresponds to the solution of the Euclidean Schrödinger equation on a fixed space-time background and in Mathematics, the heat kernel - corresponding to the determination of the spectrum of the Laplacian - can give topological information (e.g. the famous Atiyah-Singer index theorem). The following theorem proved by Minakshisundaram-Pleijel-De Witt-Gilkey gives the complete asymptotic solution for the heat kernel on a Riemannian manifold.

Theorem 2.2.1 (See [20] for a precise statement) *Let M be a Riemannian manifold without boundary. Then for each $x \in M$, there is a complete asymptotic expansion*

$$p(\tau, x|y) = \frac{\sqrt{g(y)}}{(4\pi\tau)^{\frac{n}{2}}} \sqrt{\Delta(x,y)} \mathcal{P}(y,x) e^{-\frac{\sigma(x,y)}{2\tau}} \sum_{n=1}^{\infty} a_n(x,y) \tau^n, \quad \tau \rightarrow 0 \quad (2.11)$$

- Here, $\sigma(x, y)$ is the Synge world function equal to one half of the square of geodesic distance $|x - y|_g$ between x and y for the metric g . This distance is defined as the minimizer of

$$|x - y|_g = \min_C \int_0^T \sqrt{g_{ij} \frac{dx^i}{dt} \frac{dx^j}{dt}} dt$$

and t parameterizes the curve C . The Euler-Lagrange equation gives the following geodesic differential equation which depends on the Christoffel's coefficients Γ_{jk}^i (2.7)

$$\frac{d^2 x^i}{dt^2} + \Gamma_{jk}^i \frac{dx^j}{dt} \frac{dx^k}{dt} = 0 \quad (2.12)$$

- $\Delta(x, y)$ is the so-called Van Vleck-Morette determinant

$$\Delta(x, y) = g(x)^{-\frac{1}{2}} \det \left(-\frac{\partial^2 \sigma(x, y)}{\partial x \partial y} \right) g(y)^{-\frac{1}{2}} \quad (2.13)$$

with $g(x) = \det[g_{ij}(x, x)]$

- $\mathcal{P}(y, x)$ is the parallel transport of the Abelian connection along the geodesic from the point y to x

$$\mathcal{P}(y, x) = e^{-\int_{C(y, x)} \mathcal{A}_i dx^i} \quad (2.14)$$

- The $a_i(x, y)$, called the heat kernel coefficients, are smooth sections $\Gamma(M \times M, \mathcal{L} \boxtimes \mathcal{L}^*)$. The first coefficient is simple

$$a_0(x, y) = 1, \quad \forall (x, y) \in M \times M$$

The other coefficients are more complex. However, when evaluated on the diagonal $x = y$, they depend on geometric invariants such as the scalar curvature R . The non-diagonal coefficients can then be computed as a Taylor series when x is in a neighborhood of y . The first diagonal coefficients are fairly easy to compute by hand. Recently $a_n(x, x)$ has been computed up to order $n = 8$. These formulas become exponentially more complicated as n increases. For example, the formula for a_6 has 46 terms.

2.2.2 Riemann surfaces

In our geometric framework, a SVM corresponds to a Riemann surface. In particular, the SABR corresponds to the Poincaré hyperbolic plane. This identification allows us to find an exact solution to the SABR model with $\beta = 0$ (see Section 2.4.1). Furthermore, we will derive a general asymptotic implied volatility for any stochastic volatility model. This expression only depends on the geometric objects that we have introduced before (i.e., the geodesic distance and the parallel transport). Although our formulation can handle arbitrary SVMs, we will assume that $\sigma^1(S, a) = C(S)a$ as all SVMs that have been introduced in the literature fit this class.

The metric associated to a one-factor SVM defined by (2.1) is (using (2.3)) in the coordinates $[x, y]$ (i.e., isothermal coordinates - see below)

$$ds^2 = e^{\phi(y)}(dx^2 + dy^2) \quad (2.15)$$

with the conformal factor

$$F(y) \equiv e^{\phi(y)} = \frac{2}{a(y)^2(1 - \rho^2)} \quad (2.16)$$

We have introduced the variables $q(S) = \int_{f_0}^S \frac{dx}{C(x)}$ and $\xi(a) = \int^a \frac{u}{\sigma(u)} du$ and the new coordinates

$$x = q(f) - \rho\xi(a) \quad (2.17)$$

$$y = (1 - \rho^2)^{\frac{1}{2}}\xi(a) \quad (2.18)$$

Note that this metric exhibits a Killing vector ∂_x as the conformal factor does not depend on the coordinate x . Due to this Killing vector field, the geodesic equation is completely integrable (in terms of quadrature) and we have (see [B1] Chapter 4 for details) the geodesic distance d between two points (x_1, y_1) and (x_2, y_2) in closed-form

$$d(x_1, y_1, x_2, y_2) = \int_{y_1}^{y_2} \frac{F(y)dy}{\sqrt{F(y) - C^2}} \quad (2.19)$$

with the constant $C = C(x_1, y_1, x_2, y_2)$ determined by the equation

$$x_2 - x_1 = \int_{y_1}^{y_2} \frac{C}{\sqrt{F(y) - C^2}} dy \quad (2.20)$$

From this expression, one can compute the exact short-time limit (at the zero-order) of the implied volatility (see Chapter 6 in [B1] and [A11]). As far as I know, no alternative (non-geometric) method can reproduce this result.

The scalar curvature is given by

$$R = \frac{\sigma(a)^2}{a} \left(\frac{\sigma'(a)}{\sigma(a)} - \frac{2}{a} \right) \quad (2.21)$$

Table 2.2.2 shows the conformal factors and the curvatures associated to the commonly used SVMs.

The metric associated to the SABR model is the standard hyperbolic metric on the Poincaré half-plane $\mathbb{H}^2 = \{z \in \mathbb{C} : \Im(z) \geq 0\}$ in the coordinates $[x, y]$

$$ds^2 = \frac{2}{\nu^2} \left(\frac{dx^2 + dy^2}{y^2} \right) \quad (2.22)$$

The 3/2-model corresponds to the metric on $\mathbb{C} \approx \mathbb{R}^2$. The scalar curvature is always non-negative and therefore we do not have a SVM exhibiting the metric on the two-dimensional sphere S^2 .

Name	Conformal factor	Scalar curvature
Geometric	$F(y) \sim y^{-2}$	$R = -1$
3/2-model	$F(y) \sim e^{\frac{-2y}{\sqrt{1-\rho^2}}}$	$R = 0$
SABR	$F(y) \sim y^{-2}$	$R = -1$
Heston	$F(y) \sim y^{-1}$	$R = -2a^{-2} < 0$

Table 2.1: Example of metrics for SVMs. \sim means modulo a multiplicative constant factor.

A SVM is therefore connected to a non-compact Riemann surface. The short-time behavior of the fundamental solution $p(t, x|y)$ to the heat kernel is valid when x and y are not on each other cut-locus. Surprisingly, for commonly used SVMs, the underlying manifold is a *Cartan-Hadamard manifold*, i.e., the scalar curvature is non positive for which the cut-locus is empty (see equations 2.21). Global conditions for hypoelliptic diffusions replacing the “not-in-cutlocus” requirement were recently derived in [36, 37].

Remark 2.2.2 (Heston model) *The most commonly used SVM, the Heston model, has a negative scalar curvature diverging at $a = 0$. This is a true singularity¹ and the manifold is not complete. The heat kernel expansion explained above is no more applicable as long as the volatility process can reach the singularity $a = 0$.*

The n -dimensional hyperbolic space \mathbb{H}^n appears naturally when considering a Libor market model where the Libors are driven by a common log-normal volatility process where n is the number of Libors (see [A12]).

2.2.3 Unified Asymptotic Implied Volatility [B1]

The general asymptotic implied volatility at the first order (for any (time-homogeneous) one-factor SVMs), depending implicitly on the metric g_{ij} (2.3) and the connection \mathcal{A}_i (2.5) on our Riemann surface, is given by (see Chapter 6 in [B1] for a derivation)

$$\begin{aligned} \sigma_{BS}(T, K, g_{ij}, \mathcal{A}_i) &= \frac{\ln(\frac{K}{f_0})}{\int_{f_0}^K \frac{df'}{\sqrt{2g^{SS}(c)}}} \left\{ 1 + \right. \\ &\frac{g^{SS}(c)T}{12} \left(-\frac{3}{4} \left(\frac{\partial_f g^{SS}(c)}{g^{SS}(c)} \right)^2 + \frac{\partial_f^2 g^{SS}(c)}{g^{SS}(c)} + \frac{1}{K^2} \right) + \frac{Tg^{SS'}(c)}{2g^{SS}(c)\phi''(c)} \left(\ln(\Delta g^{\mathcal{P}^2})'(c) - \frac{\phi'''(c)}{\phi''(c)} + \frac{g^{SS''}(c)}{g^{SS'}(c)} \right) \left. \right\} \\ &+ O(T^2) \end{aligned} \tag{2.23}$$

with $\phi = d^2(x, \alpha)$. Here c is the volatility a which minimizes the geodesic distance $d(x, \alpha)$ on the Riemann surface (with $x = (K, a)$). In particular, we obtained

$$\lim_{\tau \rightarrow 0} \mathbb{E}[C(S_\tau)^2 a_\tau^2 | S_\tau = S] = C(S)c^2$$

¹The Heston model behaves as an (Euclidean) black hole.

where $c = \underset{a}{\operatorname{argmin}} d((S, a), \alpha)$.

Δ is the Van Vleck-Morette determinant (2.13), g is the determinant of the metric and \mathcal{P} is the parallel gauge transport (2.14). The prime symbol $'$ indicates a derivative according to a . In the specific case of the SABR model, one obtains the exact zero-order term (see [18], Chapter 6 in [B1], [A11]):

$$\sigma_{\text{BS}}(T, f) = \frac{\ln \frac{f}{f_0}}{\operatorname{vol}(f)} + O(T)$$

with $\operatorname{vol}(f) = \frac{1}{\nu} \ln \left(\frac{q\nu + \alpha\rho + \sqrt{\alpha^2 + q^2\nu^2 + 2q\alpha\nu\rho}}{\alpha(1+\rho)} \right)$ with $q = \int_{f_0}^f \frac{dx}{C(x)}$. Note that for $\beta < 1$, we have

$$\sigma_{\text{BS}} \sim_{|K| \rightarrow \infty} \frac{\nu}{1 - \beta}$$

which is consistent with Lee's moment formula, in comparison with the original Hagan-al formula. This result is proved when $\rho = 0$ in [16].

2.3 Large-strike behavior: Gaussian estimates ♣ [B1]

In this section, we explain how large-strike behavior of implied volatility can be deduced for time-homogenous local volatility models. We conclude with some research proposal for the case of SVMs. A first step in the study of the implied volatility wings is achieved by the Lee moment formula [86], previously mentioned, which translates the behavior of the wings in the existence of higher moments for the forward. This formula states that the implied volatility is at-most linear in the moneyness $k = \ln K/f_0$ as $|k| \rightarrow \pm\infty$ with a slope depending on the existence of higher moments.

For example, when all the moments exist, we have

$$\limsup_{k \rightarrow \pm\infty} \frac{V_{\text{BS}}(\tau, k)^2}{|k|} = 0, \quad V_{\text{BS}}(\tau, k) \equiv \sigma_{\text{BS}}(\tau, k)\sqrt{\tau}$$

If we apply this result to the Black-Scholes log-normal process, the lim sup is rough as the implied volatility is flat. Moreover, the Lee moment formula is quite hard to use as we generally don't have at hand an analytical conditional probability, hence the difficulty to examine the existence of higher moments. The moment formula was recently sharpened by Benaim and Friz [53] and Gulisashvili [59] who show how the tail asymptotics of the log stock price translate *directly* to the large-strike behavior of the implied volatility. Their tail-wing formula reliably informs us when the lim sup in Lee moment formula [86] can be strengthened to a true limit. From this formula, we want to find the large strike asymptotics of the implied volatility for a time-homogeneous local volatility model using sharp Gaussian estimates for Schrödinger operators. This approach is covered in Chapter 10 in [B1].

2.3.1 Time-homogeneous local volatility model

We assume that under \mathbb{P} , the forward f follows a one-dimensional, time-homogeneous regular diffusion on an interval $I \subset \mathbb{R}_+$:

$$df_t = C(f_t)dW_t$$

with $f_{t=0} = f_0$. The so-called local volatility $C(f)$ is a positive continuous function on I . The Itô infinitesimal generator \mathcal{L} associated to the backward Kolmogorov equation for the conditional probability $p(t, f|f_0)$ is

$$\mathcal{L}p(t, f|f_0) = \frac{1}{2}C(f_0)^2\partial_{f_0}^2 p(t, f|f_0)$$

We now convert the backward Kolmogorov equation into a simpler form. For this purpose, we assume that $C'(f)$ and $C''(f)$ exist and are continuous on I in order to perform our (Liouville) transformation. By introducing the new coordinate $s = \sqrt{2} \int_{f_0}^f \frac{df'}{C(f')}$ which can be interpreted as the geodesic distance (Lamperti's transformation), one can show that the new function $P(t, s)$ defined by

$$p(t, f|f_0) = P(t, s) \frac{\sqrt{2C(f_0)}}{C(f)^{\frac{3}{2}}} \quad (2.24)$$

satisfies a (Euclidean) one-dimensional Schrödinger equation

$$(\partial_s^2 + Q(s))P(t, s) = \partial_t P(t, s) \quad (2.25)$$

The time-homogeneous potential is

$$Q(s) = \frac{1}{2}(\ln C)''(s) - \frac{1}{4}((\ln C)'(s))^2 \quad (2.26)$$

where the prime $'$ indicates a derivative according to s . Table 2.3.1 is a list of examples of potentials for a few particular time-homogeneous LV models. By applying Itô-Tanaka's formula on a call payoff $\max(f_t - K, 0)$, we obtain that a European call option $\mathcal{C}(\tau, k)$ with maturity τ and log-strike k can be rewritten as a local time formula:

$$\mathcal{C}(\tau, k) = \max(f_0 - K, 0) + \frac{\sqrt{C(K)C(f_0)}}{\sqrt{2}} \int_0^\tau P(t, s(K))dt \quad (2.27)$$

Gaussian estimates of Schrödinger semigroups

The history of *Gaussian estimates* of parabolic PDEs is quite rich starting with the works of Nash and Aronson on Gaussian estimates for Laplacian heat kernel equation. By Gaussian estimates, we mean in short that the fundamental solution $P(t, s)$ can be bounded by two Gaussian distributions

$$c_1 p_G(c_2 t, s|s_0) \leq P(t, s) \leq C_1 p_G(C_2 t, s|s_0)$$

LV Model	$C(f)$	Potential
Black-Scholes	f	$Q(s) = -\frac{1}{8}$
Quadratic	$af^2 + bf + c$	$Q(s) = -\frac{1}{8}(b^2 - 4ac)$
CEV	$f^\beta, 0 \leq \beta < 1$	$Q_{CEV}(s) = \frac{\beta(\beta-2)}{4(1-\beta)^2 s^2}$
LCEV	$f \min(f^{\beta-1}, \epsilon^{\beta-1})$ with $\epsilon > 0$	$s_\epsilon \equiv \frac{\sqrt{2}\epsilon^{1-\beta}}{(1-\beta)}$ $Q_{LCEV}(s) = Q_{CEV}(s) \forall s \geq s_\epsilon$ $Q_{LCEV}(s) = -\frac{1}{8}\epsilon^{2(\beta-1)} \forall s < s_\epsilon$

Table 2.2: Example of potentials associated to LV models.

with $p_G(t, y|x) = \frac{1}{(4\pi t)^{\frac{1}{2}}} \exp(-\frac{(y-x)^2}{4t})$ the Gaussian heat kernel and $\{C_i, c_i\}_{i=1,2}$ some constants.

The fundamental solution of (2.25) satisfies Gaussian bounds provided that the potential $Q(s)$ belongs to the Kato class:

Definition 2.3.1 (Autonomous Kato class) *We say that $Q(\cdot)$ is in the Kato class K if*

$$\limsup_{\delta \rightarrow 0} \sup_{x \in \mathbb{R}} \int_0^\delta dt \int_{\mathbb{R}} dy |Q(y)| p_G(t, y|x) = 0$$

One can check that the models listed in (Table 2.3.1) except the CEV model (due to the singularity at $s = 0$) belong to the Kato class.

Assuming that the scalar potential (2.26) associated to a separable local volatility function belongs to the Kato class, we have the Gaussian bounds on the function $P(t, s)$. This inequality directly translates on an estimation of the fundamental solution $p(t, f|f_0)$ using the relation between $p(t, f|f_0)$ and its gauge-transform $P(t, s)$ (2.24)

$$c_1 e^{c_2 t} p_G(t, s) \leq \frac{C(f)^{\frac{3}{2}}}{\sqrt{2C(f_0)}} p(t, f|f_0) \leq C_1 e^{C_2 t} p_G(t, s) \quad (2.28)$$

In the following example, we will check the (asymptotic) validity of these Gaussian estimates for the CEV model, although the potential Q does not belong to the Kato class, for which we know analytically the conditional probability (this is the case for the LCEV model).

Example 2.3.2 (CEV model) *The CEV model is a LV model for which the local volatility function is a power of the forward: $C(f) = f^\beta$, with $0 < \beta < 1$. A closed-form expression for the conditional probability is given by*

$$p(t, f|f_0) = \frac{f^{\frac{1}{2}-2\beta}}{(1-\beta)t} \sqrt{f_0} e^{-\frac{f^{2(1-\beta)} + f_0^{2(1-\beta)}}{2(1-\beta)^2 t}} \mathbf{I}_{\frac{1}{2(1-\beta)}} \left(\frac{(ff_0)^{1-\beta}}{(1-\beta)^2 t} \right)$$

where \mathbf{I} is the modified Bessel function of the first kind. As

$$\mathbf{I}_{\frac{1}{2(1-\beta)}}(x) \sim_{x \rightarrow \infty} \frac{1}{\sqrt{2\pi x}} e^x$$

we deduce that the large forward limit exhibits a Gaussian behavior

$$\frac{\sqrt{2C(f_0)}}{C(f)^{\frac{3}{2}}} p(t, f|f_0) \sim \frac{1}{\sqrt{4\pi t}} e^{-\frac{s(f)^2}{4t}} \quad (2.29)$$

with $s(f) = \int_0^f x^{-\beta} dx$. The potential associated to the CEV model is

$$Q(s) = \frac{\beta(\beta - 2)}{4(1 - \beta)^2 s^2}$$

By plugging our Gaussian estimates for the conditional probability (2.28) into the expression (2.27) and by doing the integration over the time t , we obtain lower and upper bounds for the fair value of a European call option. Then using Benaim-Fritz result [53], we can directly translate the Gaussian bounds on the call option into bounds on the implied volatility. For a time-homogeneous local volatility model, we obtain

Theorem 2.3.3 (See Chapter 10 in [B1]) *By assuming that the quantity $-\frac{1}{2} \ln(C(K)) + \frac{s(K)^2}{4\tau}$ is a regularly varying function with index $\alpha > 0$ (in k or K) and that the potential (2.26), associated to a time-homogeneous LVM, belongs to the Kato class K , the large strike behavior of the implied volatility is given by*

$$\frac{V_{\text{BS}}(\tau, k)^2}{k} \sim_{k \rightarrow \infty} \Psi \left(\frac{-\frac{1}{2} \ln(C(K)) + \frac{s(K)^2}{4\tau}}{k} \right)$$

with $\Psi(x) = 2 - 4(\sqrt{x^2 + x} - x)$ and $s(K) = \sqrt{2} \int_{f_0}^K \frac{dx}{C(x)}$. Moreover, if $s(K)$ is the leading term, we have that the large-strike behavior of the implied volatility involves the harmonic average of the local volatility function

$$\sigma_{\text{BS}}(\tau, k) \sim_{k \rightarrow \infty} \frac{\sqrt{2}k}{s(K)} \quad (2.30)$$

Remark 2.3.4 *The regularly varying assumption is not needed if we use the refined result of Gulisashvili [59]. Note that the limit (2.30) is the Berestycki-Busca-Florent formula ([17]): In the limit $\tau \rightarrow 0$, the implied volatility is the harmonic mean of the local volatility, namely*

$$\lim_{\tau \rightarrow 0} \sigma_{\text{BS}}(\tau, k) = \frac{\sqrt{2}k}{s(K)}$$

Therefore assuming that Q belongs to the Kato class, the large-strike and short-time behaviors coincide.

As an example, we obtain the following tail estimates for the implied volatility in the (L)CEV model

Example 2.3.5 ((L)CEV model) For $0 \leq \beta < 1$, we have

$$\sigma_{\text{BS}}(k, \tau) \sim_{k \rightarrow \infty} \frac{k(1-\beta)}{K^{1-\beta}} \quad (2.31)$$

and for $\beta = 1$, we have $\sigma_{\text{BS}}(\tau, k) \sim_{k \rightarrow \infty} 1$.

We have not been able to perform a similar analysis for SVMs. Gaussian estimates for Schrödinger's heat kernel equation have been obtained by assuming that the scalar curvature is bounded for below - see the famous paper by Li and Yau [88]. This assumption is not satisfied for commonly used SVM, except the SABR model for which $R = -1$ for which a similar analysis can be performed. The study of large-strike asymptotics for SVMs deserves further study. A recent breakthrough was achieved in [36, 37, 53] where the authors obtain the large-strike asymptotics of a correlated Stein-Stein model using a Varadhan-type estimate in a small-noise regime.

2.4 Solvable local and SVMs [A13]

In [A13], we classify time-homogeneous local and stochastic volatility models for which the Kolmogorov equation can be exactly solved in terms of (sum of) hypergeometric functions. One of the main examples in this classification is the SABR model with $\beta = 0$.

2.4.1 Normal SABR

For the SABR model with $\beta = 0$, the function Q and the potential \mathcal{A} vanish. p satisfies a heat kernel equation where the differential operator D reduces to the Laplace-Beltrami operator on \mathbb{H}^2 :

$$\frac{\partial p}{\partial \tau'} = \Delta_{\mathbb{H}^2} p \quad (2.32)$$

$$\equiv y^2(\partial_x^2 + \partial_y^2)p \quad (2.33)$$

with $\tau' = \frac{\nu^2 \tau}{2}$. Therefore solving the Kolmogorov equation for the normal SABR model (i.e., $\beta = 0$) is equivalent to solving this (Laplacian) heat kernel on \mathbb{H}^2 . Surprisingly, there is an analytical solution for the heat kernel on \mathbb{H}^2 (2.32) found by McKean. The exact conditional probability density p depends on the hyperbolic distance $d(z, z')$ and is given by

$$p(\tau', d) = \frac{\sqrt{2}e^{-\frac{\tau'}{4}}}{(4\pi\tau')^{\frac{3}{2}}} \int_{d(z, z')}^{\infty} \frac{be^{-\frac{b^2}{4\tau'}}}{(\cosh b - \cosh d(z, z'))^{\frac{1}{2}}} db$$

From this expression, one can derive an exact expression for call options (see Chapter 6 in [B1]). Note that swaption smiles on JPY markets can be properly calibrated with $\beta = 0$ (this was the case in 2004-2005).

2.4.2 Superpotential and local volatility

We have seen that the Kolmogorov equation for time-homogeneous local volatility model can be cast into a Schrödinger equation. A similar reduction can be performed for SVMs in which the connection \mathcal{A} is exact and therefore can be eliminated by a gauge transformation. Below, we sketch our approach in the case of a local volatility model which involves the *supersymmetrisation* of the Schrödinger equation:

$$-(\partial_s^2 + Q(s))\phi_\lambda(s) = \lambda\phi_\lambda(s) \quad (2.34)$$

where $\phi_\lambda(s)$ are eigenvectors of the self-adjoint (or self-adjoint extensions) Schrödinger operator $\partial_s^2 + Q$. Let us write (2.34) as

$$\lambda\phi_\lambda^{(1)} = A_1^\dagger A_1 \phi_\lambda^{(1)} \quad (2.35)$$

where we have introduced the first order operator A_1 and its adjoint A_1^\dagger (with respect to L^2 scalar product)

$$A_1 = \partial_s + W^{(1)}(s), \quad A_1^\dagger = -\partial_s + W^{(1)}(s)$$

$W^{(1)}$ is called a *superpotential* which satisfies the Riccati equation

$$Q^{(1)}(s) = \partial_s W^{(1)}(s) - W^{(1)}(s)^2 \quad (2.36)$$

Surprisingly, this equation is trivially solved for our specific expression for Q (even with a drift $\mu(f)$!, see [A11])

$$W^{(1)}(s) = \frac{1}{2} \frac{d \ln \sigma^{(1)}(s)}{ds} \quad (2.37)$$

Next, we define the *Scholes-Black equation* by intervening the operator A_1 and A_1^\dagger

$$\begin{aligned} \lambda\phi_\lambda^{(2)}(s) &= A_1 A_1^\dagger \phi_\lambda^{(2)}(s) \\ &= -\left(\partial_s^2 + Q^{(2)}(s)\right) \phi_\lambda^{(2)}(s) \end{aligned} \quad (2.38)$$

This corresponds to a new Schrödinger equation with the partner potential

$$Q^{(2)}(s) = -\partial_s W^{(1)} - (W^{(1)})^2 \quad (2.39)$$

By plugging our expression for the superpotential (2.37) in (2.39), we have

$$Q^{(2)} = -\frac{1}{2}(\ln \sigma^{(1)})''(s) - \frac{1}{4}(\ln \sigma^{(1)})'(s)^2 + \frac{\mu^{(1)'(s)}}{\sqrt{2}\sigma^{(1)}(s)} - \frac{\mu^{(1)}(s)^2}{2\sigma^{(1)}(s)^2} \quad (2.40)$$

Now we want to show that provided that we can solve the equation (2.35), then we have automatically a solution to (2.38) and vice versa. The SUSY-partner Hamiltonians $H_1 = A_1^\dagger A_1$ and $H_2 = A_1 A_1^\dagger$ obey the relations $A_1^\dagger H_2 = H_1 A_1^\dagger$ and $H_2 A_1 = A_1 H_1$. As a consequence H_1 and H_2

are isospectral: It means that the strictly positive eigenvalues all coincide and the corresponding eigenvectors are related by the supercharge operators A_1 and A_1^\dagger :

▷ If H_1 admits a zero eigenvalue (i.e., broken supersymmetry), we have the relation

$$\begin{aligned}\phi_0^{(1)}(s) &= C e^{-\int^s W^{(1)}(z) dz}, \quad \lambda^{(1)} = 0 \\ \phi_\lambda^{(2)}(s) &= (\lambda)^{-\frac{1}{2}} A_1 \phi_\lambda^{(1)}(s), \quad \lambda^{(2)} = \lambda^{(1)} = \lambda \neq 0 \\ \phi_\lambda^{(1)}(s) &= (\lambda)^{-\frac{1}{2}} A_1^\dagger \phi_\lambda^{(2)}(s)\end{aligned}\tag{2.41}$$

▷ If H_1 (and H_2) doesn't admit a zero eigenvalue (i.e., unbroken supersymmetry)

$$\begin{aligned}\lambda^{(2)} &= \lambda^{(1)} = \lambda \neq 0 \\ \phi_\lambda^{(2)}(s) &= (\lambda)^{-\frac{1}{2}} A_1 \phi_\lambda^{(1)}(s) \\ \phi_\lambda^{(1)}(s) &= (\lambda)^{-\frac{1}{2}} A_1^\dagger \phi_\lambda^{(2)}(s)\end{aligned}$$

It can appear that the Scholes-Black equation is much simpler to solve than the Black-Scholes PDE. This leads to a classification of solvable one and two dimensional SDEs where the conditional density can be written as a sum of hypergeometric functions (see [A11] for details). This approach using the supersymmetric structure of the Schrödinger operator leads to a large classification extending the works in [3, 30]. We illustrate this with the CEV process with $\beta = 2/3$ for which the supersymmetric partner is a Wiener process.

Example 2.4.1 (CEV with $\beta = 2/3$ and Bachelier process) *The superpotential associated with the CEV process is given by*

$$W^{(1)}(s) = \frac{\beta}{2s(1-\beta)}$$

with the flat coordinate $s = \frac{\sqrt{2}f^{1-\beta}}{(1-\beta)} \in [0, \infty)$. The potential (2.36) is

$$Q^{(1)}(s) = \frac{\beta(\beta-2)}{4(1-\beta)^2 s^2}$$

from which we deduce the partner potential (2.39)

$$Q^{(2)}(s) = \frac{\beta(2-3\beta)}{4(1-\beta)^2 s^2}$$

This partner potential corresponds to the potential of a CEV process $df = f^B dW$ with B depending on β by

$$\frac{B(B-2)}{(1-B)^2} = \frac{\beta(2-3\beta)}{(1-\beta)^2}$$

Surprisingly, we observe that for $\beta = \frac{2}{3}$, $Q^{(2)}(s)$ cancels and the corresponding partner local volatility model is the Bachelier model $df = dW$ for which the heat kernel is given by the normal distribution. The eigenvectors of the supersymmetric Hamiltonian partner $H_2 = -\partial_s^2$ to H_1 are given

by (with the absorbing boundary condition $\phi_\lambda(0) = 0$)

$$\phi_\lambda^{(2)}(s) = \frac{\sin(\sqrt{\lambda}s)}{\sqrt{4\pi\lambda^{\frac{1}{4}}}}$$

with a continuous spectrum \mathbb{R}_+ . Applying the supersymmetric transformation (2.41), we obtain the eigenvectors for the Hamiltonian $H_1 = -\partial_s^2 + \frac{2}{s^2}$ corresponding to the CEV process with $\beta = \frac{2}{3}$

$$\begin{aligned} \phi_\lambda^{(1)}(s) &= \lambda^{-\frac{1}{2}} \left(-\partial_s + \frac{1}{s} \right) \phi_\lambda^{(2)}(s) \\ &= \frac{1}{\sqrt{4\pi\lambda^{\frac{3}{4}}}} \left(-\sqrt{\lambda} \cos(\sqrt{\lambda}s) + \frac{\sin(\sqrt{\lambda}s)}{s} \right) \end{aligned}$$

By integration, we obtain the fair value for a European call option

$$\mathcal{C}(\tau, K) = \max(f_0 - K, 0) + \frac{(f_0 K)^{\frac{1}{3}}}{\sqrt{2}} \int_0^\infty d\lambda \frac{(1 - e^{-\lambda\tau})}{\lambda} \phi_\lambda^{(1)}(s) \phi_\lambda^{(1)}(s_0)$$

This expression can be integrated out and written in terms of the cumulative distribution. The fact that the CEV model with $\beta = \frac{2}{3}$ depends on the cumulative normal distribution and is therefore related to the heat kernel on \mathbb{R}_+ has been observed empirically by numerous authors. Here we have seen that it corresponds to the fact that the supersymmetric partner potential vanishes for this particular value of β .

Chapter 3

Numerical solutions of non-linear parabolic PDEs

The valuation of European options in the Black-Scholes paradigm evolves organically *linear* second order parabolic partial differential equations (PDEs). The optionality in American-style options, or market imperfections not taken into account in the Black-Scholes framework, such as illiquidity, transaction costs, delta and gamma constraints, uncertain volatilities, uncertain default intensities, different rates for lending and borrowing, and counterparty risk lead in turn to *nonlinear* parabolic PDEs. So do the calibrations of local stochastic volatility models and local correlation models to market prices of Vanilla options. Some examples of nonlinear PDEs arising in quantitative finance are shortly listed in Table 3.1. Some of them will be expanded in the next section. This chapter is extensively covered in the book “nonlinear option pricing” and in our research papers (see [B2,A1,A5,A6,A8,A9]). Our main contributions in this field are the resolution of calibration issues in mathematical finance (for local stochastic volatility models, hybrid models, local correlation models, ...) using particle method for McKean nonlinear PDEs, the use of 2-BSDEs for uncertain volatility models and the introduction of the so-called marked branching diffusions generalizing the McKean branching diffusions.

3.1 Some examples of non-linear PDEs

In this section, we describe various nonlinear PDEs that appear in finance. Some of them are connected to stochastic control problems (i.e., Hamilton-Jacobi-Bellman PDEs). Then, we will describe for each example a (new) probabilistic algorithm that can handle the curse of dimensionality of these high-dimensional PDEs.

Table 3.1: Examples of nonlinear PDEs arising in option pricing. $\Sigma(\Gamma)$ denotes the volatility in the nonlinear PDE $\partial_t u + \frac{1}{2}\Sigma(\partial_x^2 u)^2 x^2 \partial_x^2 u = 0$

Transaction costs	$\Sigma(\Gamma)^2 = \sigma^2 + \sqrt{\frac{2}{\pi} \frac{k\sigma}{\sqrt{\delta t}}} \text{sign}(\Gamma)$
Illiquid market	$\Sigma(\Gamma)^2 = \frac{\sigma^2}{(1-\epsilon x \Gamma)^2}$
American option	$\max(\partial_t u + \frac{1}{2}\sigma^2 x^2 \partial_x^2 u, g - u) = 0$
Uncertain default rate model	$\partial_t u + \frac{1}{2}\sigma^2 x^2 \partial_x^2 u + \lambda(t, u)(u^D - u) = 0$
Uncertain volatility model	$\Sigma(\Gamma)^2 = \bar{\sigma}^2 1_{\Gamma \geq 0} + \underline{\sigma}^2 1_{\Gamma < 0}$
Local stochastic volatility model	$-\partial_t p + \frac{1}{2}\partial_x^2(\sigma^2(t, x, p)p(t, x)) = 0$
Local correlation model	$-\partial_t p + \frac{1}{2}\partial_{x_1}^2(\sigma_1^2 p) + \frac{1}{2}\partial_{x_2}^2(\sigma_2^2 p) + \partial_{x_1 x_2}(\rho(t, x_1, x_2, p)\sigma_1 \sigma_2 p) = 0$
Counterparty risk	$\partial_t u + \frac{1}{2}\sigma^2 x^2 \partial_x^2 u + \lambda(t)(u^+ - u) = 0$

3.1.1 Uncertain volatility models

We model the asset S_t by a positive local Itô $(\mathcal{F}_t, \mathbb{P})$ -martingale:

$$dS_t = \sigma_t S_t dW_t$$

The volatility, which is an \mathcal{F}_t -adapted process, is unspecified for the moment. Let us then consider an option delivering some payoff F_T at maturity T which is some function of the asset path $(S_t, 0 \leq t \leq T)$. In the uncertain volatility model (in short UVM) introduced in [6, 89], the volatility is uncertain. As a minimal modeling hypothesis, we assume only that the volatility is valued in a compact interval $[\underline{\sigma}, \bar{\sigma}]$. We then define the time- t value u_t of the option as the solution to a stochastic control problem:

$$u_t = \sup_{[t, T]} \mathbb{E}^{\mathbb{P}}[F_T | \mathcal{F}_t] \quad (3.1)$$

Here $\sup_{[t, T]}$ means that the supremum is taken over all (\mathcal{F}_t) -adapted processes $(\sigma_s)_{t \leq s \leq T}$ such that for all $s \in [t, T]$, σ_s belongs to the domain $[\underline{\sigma}, \bar{\sigma}]$. The choice (3.1) can be justified by showing that u_t corresponds to the seller's super-replication price under the assumption of uncertain volatility (see [35] and our section 1.5.2 on Skorohod embedding problem which corresponds to the case $\underline{\sigma} = 0$ and $\bar{\sigma} = \infty$). For Vanilla payoffs $F_T = g(S_T)$, the HJB PDE is the so-called Black-Scholes-Barenblatt equation:

$$\begin{aligned} \partial_t u(t, s) + \frac{1}{2} s^2 \Sigma (\partial_s^2 u(t, s))^2 \partial_s^2 u(t, s) &= 0, \quad (t, s) \in [0, T) \times \mathbb{R}_+^* \\ u(T, s) &= g(s), \quad s \in \mathbb{R}_+^* \end{aligned} \quad (3.2)$$

with $\Sigma(\Gamma) = \underline{\sigma} 1_{\Gamma < 0} + \bar{\sigma} 1_{\Gamma \geq 0}$. In multi-asset case, the Hamiltonian becomes

$$H(S, \Gamma) = \frac{1}{2} \max_{(\sigma^\alpha, \rho^{\alpha\beta})_{1 \leq \alpha < \beta \leq d} \in D} \sum_{\alpha, \beta=1}^d \rho^{\alpha\beta} \sigma^\alpha \sigma^\beta S^\alpha S^\beta \Gamma^{\alpha\beta}$$

This is a fully non-linear second-order parabolic PDE.

3.1.2 Local stochastic volatility models

A local stochastic volatility model (LSVM) is defined by the following SDE for the forward f_t of maturity T

$$df_t = a_t f_t \sigma(t, f_t) dW_t \quad (3.3)$$

where a_t is a (possibly multi-factor) stochastic process. It can be seen as an extension to the Dupire local volatility model (LVM), or as an extension of a stochastic volatility model (SVM). In an SVM, one handles only a finite number of parameters (volatility-of-volatility, spot/volatility correlations, etc.). As a consequence, one is not able to perfectly calibrate to the whole implied volatility surface. In order to be able to calibrate exactly market smiles, one “decorates” the volatility of the forward with a local volatility function $\sigma(t, f)$.

From [41], this model is exactly calibrated to market smiles if and only if

$$\sigma_{\text{Dup}}(t, f)^2 = \sigma(t, f)^2 \mathbb{E}^{\mathbb{P}}[a_t^2 | f_t = f] \quad (3.4)$$

where

$$\sigma_{\text{Dup}}(t, f)^2 = \frac{\partial_t \mathcal{C}(t, f)}{\frac{1}{2} f^2 \partial_f^2 \mathcal{C}(t, f)} \quad (3.5)$$

is the Dupire local volatility that is inferred from the market Vanilla smile. $\mathcal{C}(t, f)$ is the market price (at time 0) of a call option with strike f and maturity t written on the forward of maturity T .

Once the requirement that market marginals have to be calibrated exactly has been taken into account, SDE (3.3) can be rewritten as

$$df_t = f_t \frac{\sigma_{\text{Dup}}(t, f_t)}{\sqrt{\mathbb{E}^{\mathbb{P}}[a_t^2 | f_t]}} a_t dW_t \quad (3.6)$$

The local volatility function depends on the joint pdf $p(t, f, a)$ of (f_t, a_t) :

$$\sigma(t, f, p) = \sigma_{\text{Dup}}(t, f) \sqrt{\frac{\int p(t, f, a') da'}{\int a'^2 p(t, f, a') da'}} \quad (3.7)$$

This is an example of McKean SDEs. A McKean equation for an n -dimensional process X is an SDE in which the drift and volatility depend not only on the current value X_t of the process, but also on the probability distribution \mathbb{P}_t of X_t :

$$dX_t = b(t, X_t, \mathbb{P}_t) dt + \sigma(t, X_t, \mathbb{P}_t) \cdot dW_t, \quad \mathbb{P}_t = \text{Law}(X_t), \quad X_0 \in \mathbb{R}^n \quad (3.8)$$

where W_t is a d -dimensional Brownian motion.

Example 3.1.1 (McKean-Vlasov SDEs) *The basic prototype for a McKean equation is given by the McKean-Vlasov SDE, where for $1 \leq i \leq n$ and $1 \leq j \leq d$*

$$\begin{aligned} b^i(t, x, \mathbb{P}_t) &= \int b^i(t, x, y) \mathbb{P}_t(dy) = \mathbb{E}[b^i(t, x, X_t)] \\ \sigma_j^i(t, x, \mathbb{P}_t) &= \int \sigma_j^i(t, x, y) \mathbb{P}_t(dy) = \mathbb{E}[\sigma_j^i(t, x, X_t)] \end{aligned} \quad (3.9)$$

Here the drift is just the mean value $\int b^i(t, X_t, y) \mathbb{P}_t(dy)$ of some function $b^i(t, X_t, \cdot)$ with respect to the distribution \mathbb{P}_t of X_t , and likewise for the volatility. We will also assume that $b(t, x, y)$ and $\sigma(t, x, y)$ are Lipschitz in x and y .

The probability density function $p(t, y)dy \equiv \mathbb{P}_t(dy)$ of X_t is solution to the Fokker-Planck PDE:

$$\begin{aligned} -\partial_t p(t, x) - \sum_{i=1}^n \partial_i (b^i(t, x, \mathbb{P}_t) p(t, x)) \\ + \frac{1}{2} \sum_{i,j=1}^n \partial_{ij} \left(\sum_{k=1}^d \sigma_k^i(t, x, \mathbb{P}_t) \sigma_k^j(t, x, \mathbb{P}_t) p(t, x) \right) = 0 \end{aligned} \quad (3.10)$$

with the initial condition $\lim_{t \rightarrow 0} p(t, x) = \delta(x - X_0)$ It is nonlinear because $b^i(t, x, \mathbb{P}_t)$ and $\sigma_k^i(t, x, \mathbb{P}_t)$ depend on the unknown p .

In [105, 92], uniqueness and existence are proved for equation (3.8) if the drift and volatility coefficients are Lipschitz-continuous functions of x (and a linear growth condition) and \mathbb{P}_t , with respect to the so-called Wasserstein distance (this is verified for the McKean-Vlasov SDEs). Note these conditions are not fulfilled by SDE (3.6) and therefore the uniqueness and existence result is not obvious. Our numerical tests in [B2, A6] indicates that this SDE is not well-defined (i.e., the calibration procedure is not working even with a large number of particles or with a PDE implementation) if the smile produced by the naked SVM (i.e. $\sigma(t, f) = 1$) is very different from the market smile. The mathematical analysis deserves further studies ♣. A partial result exists: in [1] it is shown that the calibration problem for a LSVM is well posed but only (a) until some maturity T^* , (b) if the volatility-of-volatility is small enough, and (c) in the case of suitably regularized initial conditions - hence the result does not apply to Equation (3.10) because of the initial Dirac mass.

3.1.3 Counterparty risk valuation

The recent financial crisis has highlighted the importance of *credit valuation adjustment* (in short CVA) when pricing derivative contracts. Bilateral counterparty risk is the risk that the issuer of a derivative contract or the counterparty may default prior to the expiry and fails to make future payments. This market imperfection leads naturally for Markovian models to semilinear PDEs. More precisely, the nonlinearity in the pricing equation affects none of the differential terms and depends on the positive part of the mark-to-market value of the derivative upon default. In short,

depending on the (modeling) choice of the mark-to-market value of the derivative upon default, we will get two types of semilinear PDEs that can be schematically written as

$$\partial_t u + \mathcal{L}u + r_0 u + r_1 u^+ = 0, \quad u(T, x) = g(x) \quad (3.11)$$

and

$$\begin{aligned} \partial_t u + \mathcal{L}u + r_0 u + r_1 M + r_2 M^+ &= 0, & u(T, x) &= g(x) \\ \partial_t M + \mathcal{L}M + r_3 M &= 0, & M(T, x) &= g(x) \end{aligned} \quad (3.12)$$

\mathcal{L} is the Itô generator of a multi-dimensional diffusion process and the r_i 's are functions of t and x .

PDE derivation

We assume that the issuer is allowed to dynamically trade d underlying assets X^1, \dots, X^d . Additionally, in order to hedge his credit risk on the counterparty name, he can trade a default risky bond, denoted P_t^C . Furthermore, the values of the underlyings are not altered by the counterparty default which is modeled by a Poisson jump process. For the sake of simplicity, we consider a constant intensity. This assumption can be easily relaxed, in particular the intensity can follow an Itô diffusion. For use below, expressions with a subscript C denote counterparty quantities. We consider the case of a long position (for the counterparty) in a single derivative whose value we denote u . In practice netting agreements apply to the global mark-to-market value of a pool of derivative positions - u would then denote the aggregate value of these derivatives. The processes $X_t = (X_t^1, \dots, X_t^d) \in \mathbb{R}_+^d$ and P_t^C satisfy under the risk-neutral measure \mathbb{P} (we assume the market model is complete)

$$\begin{aligned} \frac{dX_t}{X_t} &= rdt + \sigma(t, X_t).dW_t \\ \frac{dP_t^C}{P_t^C} &= (r + \lambda_C)dt - dJ_t^C \end{aligned}$$

with W_t a d -dimensional Brownian motion, J_t^C a jump Poisson process with intensity λ_C and r the interest rate. The no-arbitrage condition and the completeness of the market give that $e^{-rt}u(t, X_t)$ is a \mathbb{P} -martingale, characterized by

$$\partial_t u + \mathcal{L}u + \lambda_C(\tilde{u} - u) - ru = 0$$

where \mathcal{L} denotes the Itô generator of X and \tilde{u} the derivative value just after the counterparty has defaulted. At the default event, \tilde{u} is given by

$$\tilde{u} = RM^+ - M^-$$

with M the mark-to-market value of the derivative to be used in the unwinding of the position upon default and R the recovery rate. Indeed, if the mark-to-market value M is positive, meaning that M should be received from the counterparty, only a fraction RM will be received in case of

default of C (it is usual to take $R = 0.4$). If the mark-to-market value M is negative, meaning that M should be received by the counterparty, we will pay M in case of default of C .

In the case of bilateral counterparty risk, the above PDE is replaced by

$$\partial_t u + \mathcal{L}u + \lambda_C (\tilde{u}_C - u) + \lambda_B (\tilde{u}_B - u) - ru = 0$$

with

$$\tilde{u}_B = M^+ - RM^-$$

and λ_B the issuer's default intensity. For the sake of simplicity, we focus only on the unilateral counterparty risk below.

There is an ambiguity in the market about the convention for the mark-to-market value to be settled at default. There are two natural conventions : the mark-to-market of the derivative is evaluated at the time of default with provision for counterparty risk or without.

1. Provision for counterparty risk, $M = u$:

$$\partial_t u + \mathcal{L}u - (1 - R)\lambda_C u^+ - ru = 0, \quad u(T, x) = g(x) \quad (3.13)$$

This PDE is semilinear.

2. No provision for counterparty risk ($M(t, x) = e^{-r(T-t)}\mathbb{E}_{t,x}[g(X_T)]$):

$$\begin{aligned} \partial_t u + \mathcal{L}u + \lambda_C (RM^+ - M^- - u) - ru &= 0, & u(T, x) &= g(x) \\ \partial_t M + \mathcal{L}M - rM &= 0, & M(T, x) &= g(x) \end{aligned} \quad (3.14)$$

PDE (3.14) is here linear.

3.2 Second-order BSDEs and UVM [A8]

1-BSDEs, first introduced by Pardoux-Peng [97], provide a probabilistic representation of solutions of nonlinear parabolic PDEs, generalizing the Feynman-Kac formula. However, the corresponding PDEs cannot be nonlinear in the second order derivative and are therefore connected to HJB equations with no control on the diffusion term. In [32], Cheridito *et al.* provide a stochastic representation for solutions of fully nonlinear parabolic PDEs by introducing a new class of BSDEs, the so-called second order BSDEs (in short 2-BSDEs). Existence and uniqueness is studied in [103].

In [A8], we start applying 2-BSDEs to the option pricing in the UVM. Note that an alternative numerical approach relying on a parametrization of the optimal frontier is also provided. A similar study was independently performed in [47] for fully nonlinear HJB arising in the theory of portfolio optimization and a convergence result is derived. Note that the schemes proposed in [A8] and [47] are different.

From (3.3), the 2-BSDE associated to the BSB equation is

$$\begin{aligned}
dX_t^\alpha &= \hat{\sigma}^\alpha X_t^\alpha dW_t^\alpha, \quad d\langle W^\alpha, W^\beta \rangle_t = \hat{\rho}^{\alpha\beta} dt, \quad 1 \leq \alpha < \beta < d \\
dY_t &= -H(X_t, \Gamma_t) dt + \sum_{\alpha=1}^d Z_t^\alpha \diamond \hat{\sigma}^\alpha X_t^\alpha dW_t^\alpha \\
dZ_t^\alpha &= A_t^\alpha dt + \sum_{\beta=1}^d \Gamma_t^{\alpha\beta} \hat{\sigma}^\beta X_t^\beta dW_t^\beta \\
Y_T &= g(X_T)
\end{aligned} \tag{3.15}$$

\diamond denotes the Stratanovich convention. We are free to choose the diffusion $\sigma(\cdot, \cdot)$, we pick a lognormal dynamics for X with some constant volatility $\hat{\sigma}^\alpha$ and some constant correlation $\hat{\rho}^{\alpha\beta}$.

3.2.1 The algorithm

A numerical scheme for the BSB 2-BSDE reads:

Scheme UVM:

$$\begin{aligned}
X_{t_i}^\alpha &= X_0^\alpha e^{-(\hat{\sigma}^\alpha)^2 \frac{t_i}{2} + \hat{\sigma}^\alpha W_{t_i}^\alpha}, \quad \mathbb{E}[\Delta W_{t_i}^\alpha \Delta W_{t_i}^\beta] = \hat{\rho}^{\alpha\beta} \Delta t_i \\
Y_{t_n} &= g(X_{t_n}) \\
Y_{t_{i-1}} &= \mathbb{E}[Y_{t_i} | X_{t_{i-1}}] \\
&+ \left(H(X_{t_{i-1}}, \Gamma_{t_{i-1}}) - \frac{1}{2} \sum_{\alpha, \beta=1}^d \hat{\rho}^{\alpha\beta} \hat{\sigma}^\alpha \hat{\sigma}^\beta X_{t_{i-1}}^\alpha X_{t_{i-1}}^\beta \Gamma_{t_{i-1}}^{\alpha\beta} \right) \Delta t_i \\
&+ (\Delta t_i)^2 \hat{\sigma}^\alpha \hat{\sigma}^\beta X_{t_{i-1}}^\alpha X_{t_{i-1}}^\beta \Gamma_{t_{i-1}}^{\alpha\beta} \\
&= \mathbb{E} \left[Y_{t_i} \left(U_{t_i}^\alpha U_{t_i}^\beta - \Delta t_i \hat{\rho}_{\alpha\beta}^{-1} - \Delta t_i \hat{\sigma}^\alpha U_{t_i}^\alpha \delta_{\alpha\beta} \right) \middle| X_{t_{i-1}} \right]
\end{aligned} \tag{3.16}$$

with $U_{t_i}^\alpha \equiv \sum_{\beta=1}^d \hat{\rho}_{\alpha\beta}^{-1} \Delta W_{t_i}^\beta$. Compared to the scheme used in [47], we have changed the discretization for the Gamma Γ by introducing explicitly the Malliavin weight for a lognormal diffusion with volatility $\hat{\sigma}$ and correlation $\hat{\rho}$. This scheme performs better in our numerical experiments. A different conclusion was reached in [47] where their scheme was shown to perform better. Note that in [47] the payoffs tested are smooth ($g(x) = x^\eta$).

Scheme (3.16) requires computing $\frac{d(d+1)}{2} + 1$ conditional expectations at each discretization date. For this purpose, as for the valuation of American options, one can use different choices. Parametric regressions are used in [57] and Malliavin's weights in [24] in the case of 1-BSDEs.

The final meta-algorithm for pricing can be summarized by the following steps:

1. Simulate N_1 replications of X with a lognormal diffusion.

Table 3.2: Call spread valued using the BSDE approach. The true price (PDE) is $\mathcal{C}_{\text{PDE}} = 11.20$. Δ is the time step discretization and 2^{M_1} is the number of MC paths used.

Δ	M_1	12	13	14	15	16	17
1/2	Price	11.08	11.07	11.06	11.06	11.06	11.06
1/4	Price	11.01	11.12	11.06	11.07	11.11	11.11
1/8	Price	10.74	10.55	10.73	11.01	11.04	11.11

2. Apply the backward algorithm (3.16) using a regression approximation. In a high-dimensional problem, the parametric regression is the most appropriate.
3. Simulate N_2 independent replications of X using the gamma functions computed at the previous step. This last step is needed in order to obtain a low-biased result (a similar step is performed in the Longstaff-Schwartz algorithm for American options).

Call spread. (See Table 3.2) Let us test our algorithm in the case of a call spread option with payoff $(X_T - K_1)^+ - (X_T - K_2)^+$. We pick $K_1 = 90$ and $K_2 = 110$. The true price (PDE) is $\mathcal{C}_{\text{PDE}} = 11.20$ and the Black-Scholes price with the mid-volatility is $C_{\text{BS}} = 9.52$.

3.3 McKean SDEs and calibration in mathematical finance [A6,B2]

In this section, for the sake of simplicity, we consider as a toy example the calibration of a local volatility model with stochastic interest rate on market smiles:

$$dS_t = r_t S_t dt + \sigma(t, S_t) dW_t \quad (3.17)$$

where r_t is a stochastic interest rate. This model is commonly used by practitioners when pricing long-dated *power reverse dual currency notes* in foreign-exchange markets. For more than a decade, all attempts for designing Monte Carlo algorithms resulted in approximate methods: even if one used an infinite number of Monte Carlo paths, and an infinitesimal discretization time step, one would not exactly reprice the market smile. The particle algorithm, proposed in [A6] (see also [79] in the case of a local correlation model) and extensively developed in [B2], is an elegant simulation technique for McKean SDEs. We use this powerful tool to build an exact calibration algorithm: if one uses this algorithm with an infinite number of Monte Carlo paths, and an infinitesimal discretization time step, one exactly reprices the market smile.

Proposition 3.3.1 (see Chapter 9 in [B2]) *Model (3.17) is exactly calibrated to the market smile (i.e. $S_t \sim \mathbb{P}^t$ for all $t \leq T$) if and only if*

$$\sigma(t, K)^2 = \sigma_{\text{Dup}}(t, K)^2 - \frac{\mathbb{E}^{\mathbb{P}}[D_{0t}(r_t - r_t^0) 1_{S_t > K}]}{\frac{1}{2} K \partial_K^2 \mathcal{C}(t, K)} \quad (3.18)$$

for all (t, K) , with $r_t^0 = -\partial_t \ln P_{0t}$ (with P_{0t} the market price of a bond of maturity t) and

$$\sigma_{\text{Dup}}(t, K)^2 = \frac{\partial_t \mathcal{C}(t, K) + r_t^0 K \partial_K \mathcal{C}(t, K)}{\frac{1}{2} K^2 \partial_K^2 \mathcal{C}(t, K)}$$

3.3.1 Malliavin’s disintegration

We now give another expression of the contribution of stochastic interest rates to local volatility (see [B2,A6] for details). Numerical implementation of the particle algorithm using the alternative formula proves to produce a much more accurate and smooth estimation of the local volatility for far from the money strikes. As a consequence, it is very useful for extrapolation purposes. To derive this new formula, we will make use of the Malliavin calculus.

From the martingale representation theorem

$$r_t - r_t^0 = \int_0^t \sigma_r^t(s) dB_s^t, \quad r_t^0 = \mathbb{E}^{\mathbb{P}^t}[r_t]$$

with $\sigma_r^t(s)$ an adapted process. Note that, from Clark-Ocone’s formula, $\sigma_r^t(s) = \mathbb{E}_s^{\mathbb{P}^t}[D_s^{B^t} r_t]$ with $D_s^{B^t}$ the Malliavin derivative with respect to the Brownian motion B^t driving the rate r_t , and \mathbb{E}_s the conditional expectation given \mathcal{F}_s , the natural filtration of all the Brownian motions used. The application of Clark-Ocone’s formula to the process $1_{S_t > K}$ gives

$$P_{0t} \frac{\mathbb{E}^{\mathbb{P}^t}[(r_t - r_t^0) 1_{S_t > K}]}{\frac{1}{2} K \partial_K^2 \mathcal{C}(t, K)} = \frac{2}{K} \int_0^t \mathbb{E}^{\mathbb{P}^t}[\sigma_r^t(s) \cdot D_s^{B^t} S_t | S_t = K] ds \quad (3.19)$$

We call this trick a Malliavin “disintegration by parts”, because it transforms an unconditional expectation involving the Heaviside function $1_{S_t > K}$ into a conditional expectation given $S_t = K$. The Malliavin integration by parts formula goes the other way round and was used in [51, 52] to obtain probabilistic representation of Greeks. Note that the second derivative $\partial_K^2 \mathcal{C}(t, K)$ of the call option with respect to strike cancels out in the right hand side of Equation (3.19). This is fortunate as the computation of this term is sensitive to the strike interpolation/extrapolation method. Also, both $\mathbb{E}^{\mathbb{P}^t}[(r_t - r_t^0) 1_{S_t > K}]$ and $K \partial_K^2 \mathcal{C}(t, K)$ are very small for strikes K that are far from the money. Numerically, this 0/0 ratio can be problematic. There is no such problem in the right hand side of Equation (3.19), because of the Malliavin disintegration by parts. This makes the Malliavin representation of the contribution of stochastic interest rates to local volatility very useful in practice, in particular when one wants to design an accurate extrapolation of the contribution of stochastic interest rates to local volatility for strikes that are far from the money.

3.3.2 The algorithm: toy example

We consider a local volatility model where the short rate follows a Ho-Lee model, for which the volatility $\sigma_r(s) = \sigma_r$ is a constant. A bond of maturity T is given in this model by

$$P_{tT} = \frac{P_{0T}^{\text{mkt}}}{P_{0t}^{\text{mkt}}} \exp \left(\frac{\sigma_r^2 (T-t)^2 t}{2} - \sigma_r (T-t) B_t^T \right) \quad (3.20)$$

with a volatility $\sigma_P^T(t) = -\sigma_r(T-t)$ and P_{0T}^{mkt} the ($t=0$) market price of a bond of maturity T . From Proposition 3.18 and using formula (3.19), the local volatility is (see [B2,A6] for details)

$$\begin{aligned} \sigma(t, K, \mathbb{P}_t^T)^2 &= \sigma_{\text{Dup}}(t, K)^2 - 2\rho\sigma_r \frac{\mathbb{E}^{\mathbb{P}^T}[P_{tT}^{-1}V_t U_t | S_t = K]}{\mathbb{E}^{\mathbb{P}}[P_{tT}^{-1} | S_t = K]} \\ &\quad - 2\sigma_r^2 \frac{\mathbb{E}^{\mathbb{P}^T}[P_{tT}^{-1}V_t(t\Theta_t - \Lambda_t) | S_t = K]}{\mathbb{E}^{\mathbb{P}^T}[P_{tT}^{-1} | S_t = K]} \end{aligned} \quad (3.21)$$

with the tangent processes

$$\begin{aligned} \frac{dV_t}{V_t} &= S_t \partial_S \sigma(t, S_t) (dW_t^T + (\rho\sigma_P^T(t) - \sigma(t, S_t)) dt), \quad V_0 = 1 \\ U_t &= \int_0^t \frac{\sigma(s, S_s)}{V_s} ds, \quad \Theta_t = \int_0^t \frac{ds}{V_s}, \quad \Lambda_t = \int_0^t \Theta_s ds \end{aligned}$$

We have checked the accuracy of our calibration procedure on the DAX market smile (30-May-11). We have chosen $\sigma_r(s) = 6.3$ bps per day (1% per year) and set the correlation between the stock and the rate to $\rho = 40\%$. The time discretization $\Delta t = t_{k+1} - t_k$ has been set to $\Delta t = 1/100$ and we have used $N = 2^{10}$ or $N = 2^{12}$ particles. After calibrating the model using the particle algorithm, we have computed Vanilla smiles using a (quasi) Monte Carlo pricer with $N = 2^{15}$ paths and a time step of $1/250$. Figure 3.1 shows the implied volatility for the market smile (DAX, 30-May-11) and the hybrid Ho-Lee/Dupire model for maturities 4 years and 10 years. When we use the Malliavin representation, the computational time is around 4 seconds for maturities up to 10 years with $N = 2^{10}$ particles (12 seconds with $N = 2^{12}$). Our algorithm definitively outperforms a (two-dimensional) PDE implementation and can handle multi-factor models. In practise, this algorithm converges with $N = 2^{10}/2^{11}$ particles.

Note that our particle algorithm can be applied to numerous calibration issues in mathematical finance. Indeed marginals are always one-dimensional and therefore, these problems are connected to similar McKean SDEs for which the particle algorithm (eventually combined with Malliavin's disintegration) applies (see [B2] for examples).

3.4 Marked branching diffusions and CVA [A5,A1,B2]

Branching diffusions have been first introduced by McKean and Skorohod [91, 104] (see also Ikeda, Nagasawa, and Watanabe [74, 75, 76]) to give a probabilistic representation of the Kolmogorov-Petrovskii-Piskunov (KPP) PDE and more generally of semilinear PDEs of the type

$$\begin{aligned} \partial_t u + \mathcal{L}u + \beta(t) \left(\sum_{k=0}^{\infty} p_k u^k - u \right) &= 0 \quad \text{in } [0, T] \times \mathbb{R}^d \\ u(T, x) &= g(x) \quad \text{in } \mathbb{R}^d \end{aligned} \quad (3.22)$$

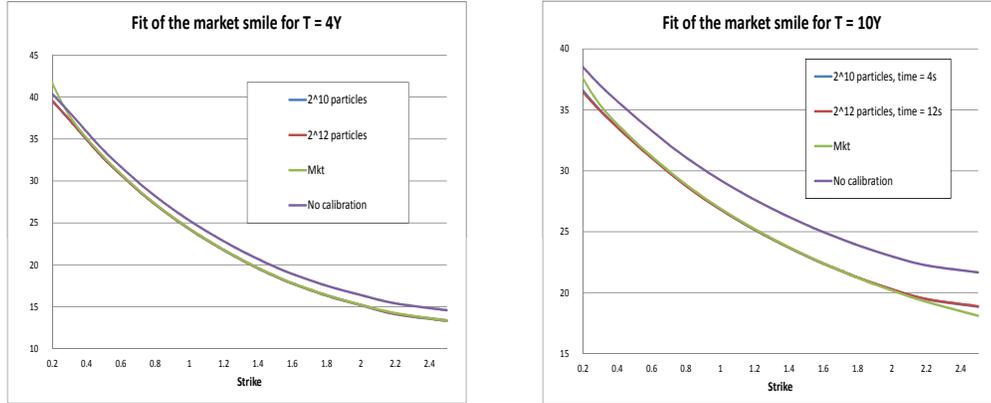


Figure 3.1: DAX (30-May-11) Implied volatilities $T = 4Y$, $T = 10Y$. Ho-Lee parameters: $\sigma_r(s) = 6.3$ bps per day, $\rho = 40\%$. Computation time with $\Delta t = 1/100$, $N = 2^{10}$ on a full 10Y implied volatility surface with a Intel Core Duo[®], 3 Ghz, 3 GB of Ram: 4s.

with $\beta \geq 0$. Here the nonlinearity is a power series in u where the coefficients p_k are required to be nonnegative and sum to one:

$$f(u) \equiv \sum_{k=0}^{\infty} p_k u^k, \quad \sum_{k=0}^{\infty} p_k = 1, \quad 0 \leq p_k \leq 1 \quad (3.23)$$

The probabilistic interpretation of such an equation goes as follows (see [74, 75, 76]). Let a single particle start at the origin, follow an Itô diffusion on \mathbb{R}^d with generator \mathcal{L} , after a mean $\beta(\cdot)$ exponential time (independent of X) dies and produces k descendants with probability p_k ($k = 0$ means that the particle dies without generating offspring). Then, the descendants perform independent Itô diffusions on \mathbb{R}^d (with the same generator \mathcal{L}) from their birth locations, die and produce descendants after a mean $\beta(\cdot)$ exponential times, etc. This process is called a d -dimensional branching diffusion with a branching rate $\beta(\cdot)$. β can also depend spatially on x or be itself stochastic (Cox process). This birth-death process describes a so-called Galton-Watson tree.

At this stage, PDE (3.22) should be compared with the semilinear PDE (3.11) arising in the pricing of counterparty risk where $f(u) = u^+$. It seems too restrictive and unreasonable to approximate the nonlinearity u^+ by a polynomial of type (3.23). A natural question is therefore to search if the probabilistic interpretation of the KPP PDE, leading to a *forward Monte Carlo scheme*, can be generalized to an arbitrary analytical nonlinearity for which the PDE is

$$\partial_t u + \mathcal{L}u + \beta(t)(F(u) - u) = 0, \quad u(T, x) = g(x) \quad (3.24)$$

with $F(u) = \sum_{k=0}^{\infty} a_k u^k$ a convergent power series in u on an interval $[-R, R]$. For example, could we cover the case $F(u) = u(1-u)(u-\epsilon) + u$ cited in [91] and not covered by branching diffusions?

For convenience, we write F as

$$F(u) = \sum_{k=0}^{\infty} \bar{a}_k p_k u^k$$

for some probabilities p_k that we leave unspecified for the moment. We have $p_k \neq 0$ and $\bar{a}_k = \frac{a_k}{p_k}$ for $a_k \neq 0$, and set $p_k = 0$ and $\bar{a}_k = 0$ otherwise. We show below that a probabilistic interpretation of (3.24) can be achieved by counting the branchings of each monomial u^k .

Assumption (Comp): In order to have uniqueness in the viscosity sense, we assume that PDE (3.24) satisfies a comparison principle for sub- and supersolutions with appropriate growth conditions.

For each Galton-Watson tree, we denote $\omega_k \in \mathbb{N}$ the number of branchings of monomial type u^k , i.e., the number of branchings where the dying particle gives birth to exactly k descendants. We call this tree endowed with the number of branchings a marked Galton-Watson tree. The descendants are drawn with an arbitrary distribution p_k (see an “optimal” choice in [B2]). We then define the multiplicative functional:

$$\hat{u}(t, x) = \mathbb{E}_{t,x} \left[\prod_{i=1}^{N_T} g(z_T^i) \prod_{k=0}^{\infty} \bar{a}_k^{\omega_k} \right], \quad \omega_k = \# \text{ branchings of type } k \quad (3.25)$$

where we recall that

$$\bar{a}_k = \begin{cases} a_k/p_k & \text{if } a_k \neq 0 \\ 0 & \text{otherwise} \end{cases}$$

and z_T^i is the position of the i -th particle at time T .

We now state the main result in [A5]:

Theorem 3.4.1 *Let us assume that $\hat{u} \in L^\infty([0, T] \times \mathbb{R}^d)$ with $\|\hat{u}\|_\infty < R$. Then the function $\hat{u}(t, x)$ is a viscosity solution to (3.24). If (Comp) holds, then $\hat{u}(t, x)$ is the unique viscosity solution.*

The requirement $\hat{u} \in L^\infty([0, T] \times \mathbb{R}^d)$ can be checked using the following proposition

Proposition 3.4.2 (see A5) *Let us assume that $g \in L^\infty(\mathbb{R}^d)$. Set $p(s) \equiv \beta(-s + \sum_{k=0}^{\infty} |a_k| \|g\|_\infty^{k-1} s^k)$.*

1. *Case $\sum_{k=0}^{\infty} |a_k| \|g\|_\infty^{k-1} > 1$: We have $\hat{u} \in L^\infty([0, T] \times \mathbb{R}^d)$ (as defined by (3.25)) if there exists $X \in \mathbb{R}_+^*$ such that*

$$\int_1^X \frac{ds}{p(s)} = T$$

In the particular case of one branching type $k \neq 1$, the sufficient condition for convergence reads as

$$|a_k| \|g\|_\infty^{k-1} \left(1 - e^{-\beta T(k-1)}\right) < 1$$

2. Case $\sum_{k=0}^\infty |a_k| \|g\|_\infty^{k-1} \leq 1$: $\hat{u} \in L^\infty([0, T] \times \mathbb{R}^d)$ for all T .

Note that, as expected, our blow-up criterion does not depend on the probabilities p_k . For $a_k = p_k$ and $\|g\|_\infty \leq 1$, we are in Case 2 and our algorithm converges as it was proved for McKean KPP branching processes.

1-BSDEs and path-dependent PDEs [A1]

Theorem 3.4.1 was generalized in the case of 1-BSDEs where PDE (3.24) is replaced by a so-called path-dependent PDE, introduced in the linear case by Dupire [43] using functional Itô calculus and extended in the nonlinear case by [45].

Let $\Omega^0 \equiv \{\omega \in C([0, T], \mathbb{R}^d) : \omega_0 = \mathbf{0}\}$ be the canonical space of continuous paths with initial value $\mathbf{0}$, \mathbb{F}^0 the canonical filtration and $\Lambda^0 \equiv [0, T] \times \Omega^0$. For every $(t, \omega) \in \Lambda^0$, denote $\|\omega\|_t \equiv \sup_{0 \leq s \leq t} |\omega(s)|$. Then the canonical process $B(\omega) = \{B_t(\omega) \equiv \omega_t, 0 \leq t \leq T\}$ for all $\omega \in \Omega^0$, defines a Brownian motion under the Wiener measure \mathbb{P}_0 .

Let $\mu : \Lambda^0 \rightarrow \mathbb{R}^d$ and $\sigma : \Lambda^0 \rightarrow \mathbb{S}^d$ be \mathbb{F}^0 -progressively measurable processes. Suppose further that for every $0 \leq t \leq t' \leq T$ and $\omega, \omega' \in \Omega^0$,

$$|\mu(t, \omega) - \mu(t', \omega')| + |\sigma(t, \omega) - \sigma(t', \omega')| \leq C(\sqrt{|t - t'|} + \|\omega_{t \wedge \cdot} - \omega'_{t' \wedge \cdot}\|_T) \quad (3.26)$$

for some constant $C > 0$, and $\sigma \sigma^T(t, \omega) \geq c_0 I_d$ for some constant $c_0 > 0$. We denote, for every $(t, \mathbf{x}) \in \Lambda^0$, by ${}^{t, \mathbf{x}}X$ the solution of the following SDE under \mathbb{P}_0 :

$$X_s = \mathbf{x}_s, \quad \forall s \leq t \quad \text{and} \quad X_s = \mathbf{x}_t + \int_t^s \mu(r, X_\cdot) dr + \int_t^s \sigma(r, X_\cdot) dB_r, \quad \forall s > t. \quad (3.27)$$

Suppose that $\psi : \Omega^0 \rightarrow \mathbb{R}$ is a non-zero, bounded Lipschitz continuous function, and $F : (t, \mathbf{x}, y) \in \Lambda^0 \times \mathbb{R} \rightarrow \mathbb{R}$ is a function Lipschitz in y such that for every y , $F(\cdot, y)$ defined on Λ^0 is \mathbb{F}^0 -progressive. We consider the following BSDE:

$$Y_t = \psi({}^{0, \mathbf{0}}X_\cdot) + \int_t^T F(s, {}^{0, \mathbf{0}}X_\cdot, Y_s) ds - \int_t^T Z_s dB_s, \quad \mathbb{P}_0 - a.s., \quad (3.28)$$

where the generator F has the following power series representation in y , locally in (t, \mathbf{x}) :

$$F(t, \mathbf{x}, y) \equiv \beta \left(\sum_{k=0}^\infty a_k(t, \mathbf{x}) y^k - y \right), \quad (t, \mathbf{x}) \in \Lambda^0, \quad (3.29)$$

for some constant $\beta > 0$, and some sequence $(a_k)_{k \geq 0}$ of bounded scalar \mathbb{F}^0 -progressive functions defined Λ^0 . We also assume that every a_k is uniformly $1/2$ -Hölder-continuous in t and Lipschitz-continuous in ω .

Denoting by $|\cdot|_0$ the $L^\infty(\Lambda^0)$ -norm, we now formulate conditions on the power series

$$\ell_0(s) \equiv \sum_{k \geq 0} |a_k|_0 s^k \quad \text{and} \quad \ell(s) \equiv \beta [|\psi|_0^{-1} \ell_0(s|\psi|_0) - s], \quad s \geq 0, \quad (3.30)$$

so as to guarantee the existence and uniqueness of a solution of the backward SDE (3.28).

Assumption .2 (i) *The power series ℓ_0 has a radius of convergence $0 < R \leq \infty$, i.e. $\ell_0(s) < \infty$ for $|s| < R$ and $\ell_0(s) = \infty$ for $|s| > R$. Moreover, the function ℓ satisfies either one of the following conditions:*

($\ell 1$) $\ell(1) \leq 0$,

($\ell 2$) *or, $\ell(1) > 0$ and for some $\hat{s} > 1$, $\ell(s) > 0, \forall s \in [1, \hat{s}]$ and $\ell(\hat{s}) = 0$.*

($\ell 3$) *or, $\ell(s) > 0, \forall s \in [1, \infty)$ and $\int_1^{\hat{s}} \frac{1}{\ell(s)} ds = T$, for some constant $s \in (1, \frac{R}{|\psi|_0})$.*

(ii) *The terminal function satisfies $|\psi|_0 < R$.*

Under Assumption .2, one proves in Theorem 2.13, [A1] that the solution (Y_t, Z_t) to the BSDE can be written in terms of a marked branching diffusion. The proof uses the notion of viscosity solutions for a path-dependent PDE.

3.4.1 Algorithm

The algorithm for solving (3.24) or (3.28) can be described by the following steps:

1. Simulate the processes X_t (with infinitesimal generator \mathcal{L}) and the Poisson default time with intensity β . Note that the intensity β can be stochastic (Cox process).
2. At each default time, produce k descendants with probability p_k (see [B2] for an “optimal” choice).
3. Evaluate for the particles alive the payoff

$$\prod_{i=1}^{N_T} g(z_T^i) \prod_{k=0}^M \left(\frac{a_k}{p_k} \right)^{\omega_k}$$

where ω_k denotes the number of branchings of type k .

In this example, we consider the degenerate PDEs:

$$\partial_t v_1 + x \partial_a v_1 + \frac{1}{2} \sigma^2 x^2 \partial_x^2 v_1 + \beta(v_1^2 - v_1) = 0, \quad \text{PDE1} \quad (3.31)$$

$$\partial_t v_2 + x \partial_a v_2 + \frac{1}{2} \sigma^2 x^2 \partial_x^2 v_2 + \beta(-v_2^2 - v_2) = 0, \quad \text{PDE2} \quad (3.32)$$

N	Fair (PDE1)	Stdev (PDE1)	Fair (PDE2)	Stdev (PDE2)
12	5.69	0.16	5.36	0.16
14	5.61	0.08	5.23	0.08
16	5.50	0.04	5.15	0.04
18	5.52	0.02	5.16	0.02
20	5.53	0.01	5.16	0.01
22	5.54	0.00	5.17	0.01

Table 3.3: MC price quoted in percent as a function of the number of MC paths 2^N . PDE pricer (PDE1) = **5.54**. PDE pricer (PDE2) = **5.17** (CPU PDE: 10 seconds). $T = 2$ years. Nonlinearity for PDE1 (resp. PDE2): $F_1(u) = u^2$ (resp. $F_2(u) = -u^2$). For completeness, the price with $\beta = 0$ (which can be obtained using a classical Monte Carlo pricer) is 6.52.

with $v_1(T, x, a) = v_2(T, x, a) = g(x, a) = (a - 1)^+$. These PDEs correspond to the backward SDEs

$$dX_t = \sigma X_t dB_t, \quad X_0 = 1 \quad (3.33)$$

$$dA_t = X_t dt, \quad A_0 = 0 \quad (3.34)$$

$$dY_t = -\beta(F(Y_t) - Y_t) dt + Z_t dB_t, \quad Y_T = g(X_T, A_T) \quad (3.35)$$

with the nonlinearities $F_1(y) = y^2$ and $F_2(y) = -y^2$. In our numerical experiments, we have taken a diffusion coefficient $\sigma = 0.2$ and a Poisson intensity $\beta = 0.1$, and the maturity $T = 2$ or $T = 5$ years. For $T = 2$ years (resp. 5 years), the probability of default is around 0.18 (resp. 0.39).

Our branching diffusion algorithm has been checked against a two-dimensional PDE solver with an ADI scheme (see Tables 3.3, 3.4). The degenerate PDEs have been converted into elliptic PDEs by introducing the process $\tilde{A}_t = \int_0^t X_s ds + (T - t)X_t$, satisfying $d\tilde{A}_t = (T - t)dX_t$.

Note that our algorithm converges to the exact PDE result as expected and the error is properly indicated by the Monte Carlo standard deviation estimator (see column Stdev). In order to illustrate the impact of the non-linearity F on the price v , we have also reported the price corresponding to $\beta = 0$. The application of this forward Monte-Carlo scheme to CVA requires a polynomial approximation of u^+ . See [B2], Section 3.4.6 for an automatic construction of such a polynomial approximation.

3.5 Marked Malliavin branching diffusions ♣ [P4]

What is appealing with the marked branching algorithm is that we get a *forward Monte-Carlo scheme*. In particular our algorithm can take care of path-dependent payoffs - this is hardly impossible with deterministic methods. As being forward and not backward as in the case of the BSDE algorithm, this scheme does not require computing conditional expectations which are difficult, specially in high-dimensional. The BSDE approach still suffers from the curse of dimensionality, a BSDE scheme can be seen as a finite element method. This is not the case of our branching algorithm. We conclude this chapter with a proposal for an extension of our marked

N	Fair (PDE1)	Stdev (PDE1)	Fair (PDE2)	Stdev (PDE2)
12	7.40	0.25	5.63	0.26
14	7.28	0.12	5.60	0.13
16	7.20	0.06	5.47	0.07
18	7.24	0.03	5.48	0.03
20	7.24	0.02	5.50	0.02
22	7.24	0.01	5.51	0.01

Table 3.4: MC price quoted in percent as a function of the number of MC paths 2^N . PDE pricer (PDE1) = **7.24**. PDE pricer (PDE2) = **5.51** (CPU PDE: 25 seconds). $T = 5$ years. Nonlinearity for PDE1 (resp. PDE2): $F_1(u) = u^2$ (resp. $F_2(u) = -u^2$). For completeness, the price with $\beta = 0$ (which can be obtained using a classical Monte Carlo pricer) is 10.24.

branching diffusions to fully non linear PDEs where the branching diffusions will be weighted by some Malliavin's weights.

Before considering the fully nonlinear case, we explain the flavor of this algorithm on (one-dimensional) linear PDEs.

Linear PDEs

Consider a one-dimensional SDE which can be converted using Lamperti's transformation into

$$dX_t = A(X_t)dt + dB_t$$

The function defined by $u(t, x) = e^{-\beta(T-t)}\mathbb{E}[\psi(X_T)|X_t = x]$ is solution of

$$\partial_t u + \frac{1}{2}\partial_x^2 u + \beta(\alpha(x)\partial_x u - u) = 0, \quad v(T, x) = \psi(x)$$

with $\alpha(x) \equiv \frac{A(x)}{\beta}$. We see this linear PDE as an heat kernel equation with a (non)linearity $\alpha(x)\partial_x v$. If the (non)linearity was $\alpha(x)v$, we can use the branching diffusion: we simulate a Poisson intensity β and at each default, we generate a single descendant with weight $\alpha(X_\tau)$. With the derivative term $\alpha(x)\partial_x v$, the above algorithm is modified by introducing a weight given by the Malliavin's weight for the first order derivative: $\mathbb{E}[\psi(X_T)|X_t = x]$ can be represented as

$$e^{-\beta(T-t)}\mathbb{E}[\psi(X_T)|X_t = x] = \mathbb{E}[F_T|W_t = x], \quad F_T \equiv \psi(W_T) \prod_{i=1}^{\#Default} \alpha(W_{\tau_i}) \frac{W_{\tau_i} - W_{\tau_{i-1}}}{\tau_i - \tau_{i-1}} \quad (3.36)$$

where $(W_{\tau_i} - W_{\tau_{i-1}})/(\tau_i - \tau_{i-1})$ corresponds to the Malliavin weight for the delta. This representation gives an exact discretization scheme for a one-dimensional SDE as it requires only the (exact) simulation of a Brownian motion and a constant Poisson's intensity. Note that exact an simulation of one-dimensional SDEs has been first considered in [21] but uses a rejection sampling for simulating the Girsanov weight.

Renormalization

From a numerical point of view, note that the functional F_T is singular in the region where $\tau_{i+1} - \tau_i \rightarrow 0$ and the variance of our algorithm diverges. We suggest to cure this instability using a *renormalization* procedure. For example, the term $\psi(W_{t_1} + \Delta W_1) \frac{\Delta W_1}{\Delta t_1}$, which is singular when $\Delta t_1 \rightarrow 0$, can be replaced by

$$(\psi(W_{t_1} + \Delta W_1) - \psi(W_{t_1})) \frac{\Delta W_1}{\Delta t_1}$$

where the added counterterm $\psi(W_{t_1}) \frac{\Delta W_1}{\Delta t_1}$ has zero mean value. If $\psi \in C_b^1$ then $|\psi(W_{t_1} + \Delta W_1) - \psi(W_{t_1})| \leq K|\Delta W_1|$ and the singularity disappears.

The overall renormalization scheme (eliminating automatically the singularity when $\tau_{i+1} - \tau_i \rightarrow 0$) can be obtained by introducing additional fictive particles: First, at each default τ_i , we create with probability 1 an independent particle, denoted A, following a Brownian motion up to time τ_i as explained previously. Second, we create a new Galton-Watson tree which is the concatenation of the previous one at time τ_i with a fictive particle, created at τ_i , which follows the Brownian motion of particle A $W_{\tau_{i+1}} - W_{\tau_i}$ weighted by a null volatility. Each Galton-Watson tree follows then this construction. Besides the global weight associated to each Galton-Watson tree, each tree is also weighted by $(-1)^{\text{Number of fictive particles created}}$. In Figure 3.2, we have plotted all Galton-Watson trees up to 2 default with the associated counterterms. These trees correspond to the expansion:

$$\begin{aligned} & \mathbb{E}[\psi(x + W_1)1_{\tau > T}] + \mathbb{E}[\{(\psi(x + W_T) - \psi(x + W_{\tau_1}))\alpha(x + W_{\tau_1})\} \frac{W_T - W_{\tau_1}}{T - \tau_1} 1_{\tau_1 < T}] + \\ & \mathbb{E}[\{((\psi(x + W_T) - \psi(x + W_{\tau_2}))\alpha(x + W_{\tau_1})\alpha(x + W_{\tau_2}) + \\ & (-\psi(x + W_{\tau_1} + W_T - W_{\tau_2}) + \psi(x + W_{\tau_1}))\alpha(x + W_{\tau_1})^2)\} \frac{W_{\tau_2} - W_{\tau_1}}{\tau_2 - \tau_1} \frac{W_T - W_{\tau_2}}{T - \tau_2} 1_{\tau_1 < T} 1_{\tau_2 < T}] \\ & + \dots \end{aligned} \tag{3.37}$$

Note that by construction the terms in brackets cancel when the Brownian increments $W_T - W_{\tau_2}$ and/or $W_{\tau_2} - W_{\tau_1}$ cancel.

A toy semi-linear PDE

We now consider a KPP-type PDE where the nonlinearity $f(\partial_x u)$ depends on the first order derivative $\partial_x u$:

$$\partial_t u + \frac{1}{2} \partial_x^2 u + \beta (f(\partial_x u) - u) = 0, \quad u(T, x) = \psi(x)$$

with $f(x) = \sum_k a_k x^k = \sum_k \frac{a_k}{p_k} p_k x^k$. This PDE can be tackled by (a) doing a Fourier expansion in space and then interpreting the non linear integral equation satisfied by the Fourier modes in terms of (marked) branching process. This approach was used in [87] to study the incompressible Navier-Stokes equation in \mathbb{R}^3 . New classes where global existence and uniqueness can be proven are derived from this representation.

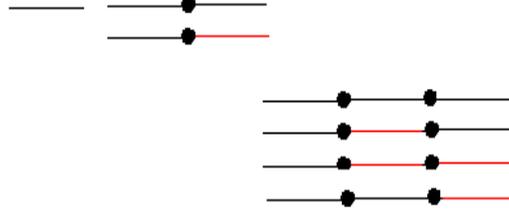


Figure 3.2: Galton-Watson trees with its renormalization up to 2 defaults. For each Galton-Watson tree, each line except the first line is weighted by the Malliavin weight and each dot is weighted by $\alpha(z)$ with z the position of the particle at this point. The red line correspond to the simulation with a zero volatility.

Example 3.5.1 (Burgers equation) We consider the Burgers equation, $f(\partial_x u) = -u\partial_x u + u$. This can be seen as a toy model for the Navier-Stokes PDE. We impose periodic boundary condition and let $u(t, x) = \sum_k \hat{u}(t, k)e^{ikx}$. Then, we obtain that each mode $\hat{u}(t, k)$ satisfies

$$\frac{d\hat{u}(t, k)}{dt} = k^2 \hat{u}(t, k) + \frac{i}{2} \hat{u} \otimes \hat{u}(t, k)$$

where \otimes denotes a convolution product with respect to a specific kernel. The numerical solution of the Burgers equation can be obtained by simulating this infinite number of particles indexed by their mode. Each particle moves (deterministically) with a drift k^2 and at each default (modeled by a Poisson intensity = 1), each particle with a mode k dies and gives birth to two new particles with modes k_1 and k_2 such that the impulsion is conserved: $k = k_1 + k_2$.

This representation of nonlinear PDEs in terms of branching diffusions in the Fourier space still suffers from the curse of dimensionality as we need to simulate an infinite number of modes. It is thus natural to search a real space branching representations. This question is open for the Navier-Stokes equation. Following the above subsection, u can be represented stochastically - written schematically

$$u(t, x) = \mathbb{E}_{t,x} \left[\prod_{i=1}^{N_T} \psi(X_T^i) \prod_{k=0} \left(\frac{a_k}{p_k} \right)^{\omega_k} \prod \text{MalliavinWeightDelta} \right]$$

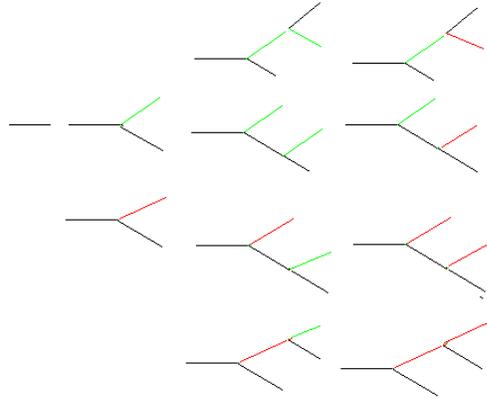


Figure 3.3: Galton-Watson trees with its renormalization up to 2 defaults for the Burgers equation - Here $f(u) = u\partial_x u$. Green lines are weighted by a Malliavin weight. The red line, weighted by a Malliavin weight, corresponds to a particle driven with a null volatility (renormalization scheme).

where X is a stochastic process with infinitesimal generator \mathcal{L} and $\text{MalliavinWeightDelta}$ denotes the Malliavin weight for the delta (for \mathcal{L}) between τ_i and τ_{i+1} . Due to the singularity $\tau_i - \tau_{i-1} \rightarrow 0$, this functional needs to be renormalized by introducing counterterms with zero mean. Following our previous discussion, each Galton-Watson tree should be normalized. For illustration purpose, let us take $f(u) = u\partial_x u$. The interaction $u\partial_x u$ means that at each default, we will generate two particles - one of type 0 and one of type 1 that will create a Malliavin weight. Each time a particle of type 1 is generated, a new Galton-Watson tree which is the concatenation of the previous one and the particle of type 1 is replaced by a fictive particle driven by a null volatility. In Figure 3.3, we have plotted all Galton-Watson trees up to two defaults.

The analysis of this algorithm is a work in progress with X. Tan and N. Touzi [P4]. We hope that it can lead to a robust forward Monte-Carlo for general semilinear parabolic second-order PDE. The renormalization procedure for the fully nonlinear case is much more involved and we wish to address it in our future research.

Publications

Books

- (B1) Henry-Labordère, P. : *Analysis, Geometry and Modeling in Finance: Advanced Methods in Option Pricing*, Financial Mathematics Series CRC, Chapman Hall (380 p.), chief-editor: M.A.H. Dempster, D. Madan, R. Cont (2008).
- (B2) Guyon, J., Henry-Labordère, P. : *Nonlinear Option Pricing*, Financial Mathematics Series CRC, Chapman Hall (440 p.), chief-editor: M.A.H. Dempster, D. Madan, R. Cont (2013).

Articles published

- (A1) Henry-Labordère, P., Touzi, N., Tan, X. : *A numerical algorithm for a class of BSDE via branching process*, accepted in *Stochastic Processes and their Applications* (2013).
- (A2) Henry-Labordère, P. : *Automated Option Pricing: Numerical Method*, accepted in *International Journal of Theoretical and Applied Finance* (2011).
- (A3) Galichon, A., Henry-Labordère, P., Touzi, N. : *A stochastic control approach to no-arbitrage bounds given marginals, with an application to Lookback options*, *Annals of Applied Probability* (2013).
- (A4) Beiglböck, M., Henry-Labordère, P., Penkner, F. : *Model-independent Bounds for Option Prices: A Mass-Transport Approach*, *Finance and Stochastics* (2013).
- (A5) Henry-Labordère, P. : *Counterparty Risk Valuation: A Marked Branching Diffusion Approach*, *Risk magazine* (Jul. 2012).
- (A6) Guyon, J., Henry-Labordère, P. : *Being particular about smile calibration*, *Risk Magazine* (Jan. 2012).
- (A7) Guyon, J., Henry-Labordère: *From Spot Volatilities to Implied Volatilities*, *Risk Magazine* (Jun. 2011).

- (A8) Guyon, J., Henry-Labordère, P. : *Uncertain Volatility Model: A Monte-Approach based on BSDE*, Journal of Computational Finance, Volume 14/Number 3, Spring 2011.
- (A9) Henry-Labordère, P. : *Calibration of Local Stochastic Volatility Models: A Monte-Approach*, Risk Magazine (Sept. 2009).
- (A10) Henry-Labordère, P. : *Short-time Asymptotics of Implied Volatility*, Encyclopedia of Quantitative Finance, Wiley (2008).
- (A11) Henry-Labordère, P. : *A Geometric Approach To Asymptotics of Implied Volatility*, Frontiers in Finance, Wiley (2008).
- (A12) Henry-Labordère, P. : *Combining the SABR and LMM models*, Risk Magazine (2007).
- (A13) Henry-Labordère, P. : *Classification of Solvable Local and Stochastic Volatility Models: Supersymmetry Methods in Option Pricing*, Quantitative Finance, Vol. 7., No. 5., 525-535 (2007).

Articles submitted

- (S1) Henry-Labordère, P. : *Linking Vanillas and VIX options: A constrained martingale optimal transport*, submitted to SIAM Journal on Financial Mathematics (2013).
- (S2) Henry-Labordère, P., Oblój, J., Spoida, P., Touzi, N. : *Maximum Maximum of Martingales given Marginals*, submitted to Annals of Applied Probability, minor revision (2013).
- (S3) Henry-Labordère, P., Touzi, N., Tan, X. : *An Explicit Martingale Version of Brenier's Theorem with infinitely many marginals*, submitted to Stochastic Processes and their Applications (2014).
- (S4) Henry-Labordère, P., Touzi, N. : *An Explicit Martingale Version of Brenier's Theorem*, submitted to Finance and stochastics (2013).
- (S5) Henry-Labordère, P. : *Vega Decomposition of Exotic Options: A Monte Carlo Approach*, SIAM Journal on Financial Mathematics, minor revision (2014).

Preprints

- (P1) Beiglböck, M., Henry-Labordère, P., Touzi, N. : *On monotone martingale transport plans*, preprint (2014).
- (P2) Henry-Labordère, P., Touzi, N.: *Martingale optimal transport*, lecture notes.
- (P3) Henry-Labordère, P., Royer, G., Touzi, N.: *Robust utility indifference pricing*, preprint (2013).

- (P4) Henry-Labordère, P., Touzi, N., Tan, X. : *Fully Non-linear PDEs: A Marked Branching Diffusion Approach*, preprint (2013).
- (P5) Henry-Labordère, P. : *Pricing in illiquid markets: From a microscopic to a macroscopic scale*, preprint (2013).

Publications in Theoretical Physics (PhD)

- (PhD1) Henry-Labordère, P., Julia, B., Paulot, L. : *Real Borcherds superalgebras and M theory*, Journal of High Energy Physics 0304:060, (2003).
- (PhD2) Henry-Labordère, P., Julia, B., Paulot, L. : *Symmetries in M theory: Monsters, Inc.* Published in Cargese (2002), Progress in string, field and particle theory (409-412).
- (PhD3) Henry-Labordère, P., Julia, B., Paulot, L. : *Borcherds symmetries in M theory*, Journal of High Energy Physics 0204:049, (2002).
- (PhD4) Henry-Labordère, P., Julia, B. : *Gravitational couplings of orientifold planes*, Journal of High Energy Physics 0201:033, (2002).
- (PhD5) Freedman, D.Z, Henry-Labordère, P. : *Field theory insight from the AdS / CFT correspondence*, Contributed to 4th Annual European TMR Conference on Integrability Nonperturbative Effects and Symmetry in Quantum Field Theory, Paris (2000).

Thesis (PhD)

Henry-Labordère, P. : *Symétries en théorie M*, PhD thesis, Ecole Normale Supérieure (Paris), 2003.

Bibliography

- [1] Abergel, F., Tachet, R. : *A nonlinear partial integrodifferential equation from mathematical finance*, Discrete Cont. Dynamical Systems, Serie A, vol. 27, 3, pp 907-917, 2010.
- [2] Acciaio, B., Beiglböck, M., Penkner, F., Schachermayer, W., Temme, J. : *A trajectorial interpretation of Doob's martingale inequalities*, Ann. Appl. Probab., To appear.
- [3] Albanese, C., Kuznetsov, A. : *Transformations of Markov Processes and Classification Scheme for Solvable Driftless Diffusions*
- [4] Albin, J. M. P.: *A continuous non-Brownian motion martingale with Brownian motion martingale distributions*, Statistics and Probability Letters, 78, 6, 682686. 2008.
- [5] Acciaio, B., Beiglböck, M., Penkner, F., Schachermayer, W. : *A Model-free Version of the Fundamental Theorem of Asset Pricing and the Super-Replication Theorem*, Preprint 2013.
- [6] Avellaneda, M., Levy, A., Paras, A. : *Pricing and Hedging Derivative Securities in Markets with Uncertain Volatility*, Applied Mathematical Finance, 1995.
- [7] Avellaneda, M., Holmes, R., Samperi, D. : *Calibrating Volatility Surfaces via Relative Entropy Minimization*, IJTAF, 1998, Proc. Courant Seminar, Vol I, 1999.
- [8] Avellaneda, M. : *Minimum-Entropy Calibration of Asset-Pricing Models*, Int. J. Theor. App. Finance 1(4) (1998) 447.
- [9] Avellaneda, M., Buff, R., Friedman, C., Grandchamp, N., Kruk, L., Newman, J. : *Weighted Monte Carlo: A New Technique for Calibrating Asset-Pricing Models*, Int. J. Theor. App. Finance, 2001, 4, 1–29.
- [10] Avellaneda, M., Boyer-Olson, D., Busca, J., Friz, P.: *Reconstruction of Volatility: Pricing index options using the steepest-descent approximation*, Risk magazine, October 2002.
- [11] Azéma, J. and Yor, M. (1979). *Une solution simple au problème de Skorokhod*, Séminaire de Probabilités XIII: 06, 90-115, LNM 721.
- [12] Beiglböck, M., Juillet, N. : *On a problem of optimal transport under marginal martingale constraints*, preprint (Aug. 2012), arXiv:1208.1509.

- [13] Beiglböck, M., Siorpaes, P. : *Pathwise Versions of the Burkholder-Davis-Gundy Inequalities*, preprint (2012).
- [14] Beiglböck, M., Huesmann, M. : *Optimal Transport and Skorokhod Embedding*, preprint (2013).
- [15] Benaim, S., Friz, P.: *Regular variation and smile asymptotics*, *Mathematical Finance*, 19(1):1–12, January (2009).
- [16] Benaim, S., Friz, P., Lee, R.: *On Black-Scholes implied volatility at extreme strikes*, *Frontiers in Quantitative Finance: Volatility and Credit Risk Modeling*, 19–45 (2008).
- [17] Berestycki, H., Busca, J., Florent, I. : *Asymptotics and Calibration of Local Volatility Models*, *Quantitative Finance*, 2:61-69 (2002).
- [18] Berestycki, H., Busca, J., Florent, I. : *Computing the Implied Volatility in Stochastic Volatility Models*, *Comm. Pure Appl. Math.*, 57, num. 10 (2004), p. 1352-1373.
- [19] Bergomi, L. : *Smile dynamics II*, *Risk Magazine*, (Oct. 2005).
- [20] Berline, N., Getzler, E., Vergne, M. : *Heat Kernels and Dirac Operators*, Springer, 1992.
- [21] Beskos, A., Roberts, G. O. : *Exact simulation of diffusions*, *Ann. Appl. Probab.* Volume 15, Number 4 (2005), 2422-2444.
- [22] Breeden, D.T. and Litzenberger, R.H. : *Prices of state-contingent claims implicit in options prices*, *J. Business*, 51, 621651 (1978).
- [23] Brown, H., Hobson, D., Rogers, L.C.G. : *Robust hedging of barrier options*, *Math. Finance*, 11(3):285314, 2001.
- [24] Bouchard, B., Touzi, N. : *Discrete-time approximation and Monte Carlo simulation of backward stochastic differential equations*, *Stochastic Processes and their Applications*, 111, 175-206 (2004).
- [25] Bouchard, B., Elie, R., Touzi, N. : *Stochastic target problems with controlled loss*, *SIAM Journal on Control and Optimization*, 48(5):3123–3150, 2009.
- [26] Bouchard, B., Nutz, M. : *Arbitrage and duality in nondominated discrete-time models*, arXiv preprint arXiv:1208.1509, 2012.
- [27] Burkholder, D. : *The best constant in the Davis inequality for the expectation of the martingale square function*, *Trans. Amer. Soc.* 354 (2002)
- [28] Carlier, G. : *On a class of multidimensional optimal transportation problems*, *Journal of convex analysis*, Vol. 10 (2003), No. 2, 517-529.
- [29] Carlier, G., Jimenez, C., Santambrogio, F. : *Optimal transportation with traffic congestion and Wardrop equilibria*, *SIAM Journal on Control and Optimization*, 47(3):1330–1350, 2008.

- [30] Carr, P., Laurence, P., Wang, T-H.: *Generating integrable one dimensional driftless diffusions*, Comptes Rendus de l'Academie des Sciences, Sept. 2006.
- [31] Carr, P., Geman, H., Madan, D.B., Yor, M. *From local volatility to local Lévy models*, Quantitative Finance, Oct. 2004.
- [32] Cheridito, P., Soner, M., Touzi, N., Victoir, N. : *Second Order Backward Stochastic Differential Equations and Fully Non-Linear Parabolic PDEs*, Communications on Pure and Applied Mathematics, Volume 60, Issue 7, July 2007, pp. 1081-1110.
- [33] Davis, M., Oblój, J., Raval, V. : *Arbitrage Bounds for Weighted Variance Swap Prices*, Mathematical Finance, 2012.
- [34] Delbaen, F., Grandits, P., Rheinlander, T., Samperi, D., Schweizer, M., Stricker, C. : *Exponential Hedging and Entropic Penalties*, Mathematical Finance, Volume 12, Issue 2, pages 99-123, April 2002.
- [35] Denis, L., Martini, C. : *A theoretical framework for the pricing of contingent claims in the presence of model uncertainty*, Ann. Appl. Probab., 16(2): 475–1058, 2006.
- [36] Deuschel, J. D., Friz, P. K., Jacquier, A., Violante, S. : *Marginal Density Expansions for Diffusions and Stochastic Volatility I: Theoretical Foundations*, Communications on Pure and Applied Mathematics, 67(1):40-82, January 2014.
- [37] Deuschel, J. D., Friz, P. K., Jacquier, A., Violante, S. : *Marginal density expansions for diffusions and stochastic volatility, part II: Applications [to the Stein–Stein model]*, preprint arXiv:1305.6765 (May 2013).
- [38] Dolinsky, Y., Soner, H.M. : *Robust hedging and martingale optimal transport in continuous time*, submitted to Probability Theory and Related Fields, (2013), forthcoming.
- [39] Dolinsky, Y., Soner, H.M. : *Robust Hedging under Proportional Transaction Costs*, Finance and Stochastics, (2013), forthcoming.
- [40] Dubins, L.E., Schwarz, G. : *On extremal martingale distributions*, Proc. Fifth Berkeley Symp. on Math. Statist. and Prob., Vol. 2, Pt. 1 (Univ. of Calif. Press, 1967), 295-299.
- [41] Dupire, B. : *Pricing with a Smile*, Risk Magazine 7, 18-20 (1994).
- [42] Dupire, B. : *A Unified Theory of Volatility*, In Derivatives Pricing: The Classic Collection, edited by Peter Carr, Risk publications.
- [43] Dupire, B. : *Functional Itô calculus*, SSRN, 2009, <http://ssrn.com/abstract=1435551>.
- [44] Dupire, B. : *Applications of the Root Solution of the Skorohod Embedding Problem in Finance*, Bloomberg (2007).
- [45] Ekren, I., Keller, C., Touzi, N., Zhang, J. : *On Viscosity Solutions of Path Dependent PDEs*, Anna. Proba. to appear.

- [46] Rouge, R., El Karoui, N. : *Pricing Via Utility Maximization and Entropy*, Mathematical Finance, Volume 10, Issue 2, pages 259-276, April 2000.
- [47] Fahim, A., Touzi, N., Warin, X. : *A probabilistic numerical method for fully nonlinear parabolic PDEs*, Ann. Appl. Probab. Volume 21, Number 4 (2011), 1322-1364.
- [48] Fan, J.Y. Hamza, K., Klebaner, F.C. : *Mimicking self-similar processes*. Preprint, 2012.
- [49] Föllmer, H : *calcul d'Itô sans probabilités*, Séminaire de Probabilités, XV: 09, 143-150, LNM 850 (1981).
- [50] Forde, M., Jacquier, A.: *The small-time smile and term structure of implied volatility under the Heston model*, SIAM Journal on Financial Mathematics, 3(1):690-708, 2012.
- [51] Fournié, E., Lasry, F. , Touzi, N. : *Monte Carlo Methods in Stochastic Volatility Models, in Numerical Methods in Financial Mathematics*, Cambridge University Press (1997).
- [52] Fournié, E., Lasry, F. , Lebuchoux, J., Lions, J. : *Applications of Malliavin Calculus to Monte Carlo Methods in Finance. II*, Finance and Stochastics, 5, 201-236 (2001).
- [53] Friz, P., De Marco, S. : *Varadhan's formula, conditioned diffusions, and local volatilities*, arXiv:1311.1545, preprint (2013).
- [54] Gangbo, W., Świąch, A. : *Optimal maps for the multidimensional Monge-Kantorovich problem*, Comm. Pure Appl. Math., 51 (1) : 23-45, 1998.
- [55] Gangbo, W., McCann, R. : *The geometry of optimal transportation*, Acta. Math. 177(2):113-161, 2006.
- [56] Gatheral, J., Hsu, E.P., Laurence, P., Ouyang, C., Wang, T-H. : *Asymptotics of implied volatility in local volatility models*, Mathematical Finance, 22(4):591-620, Oct. 2012.
- [57] Gobet, E., Lemor, J.P., Warin, X. : *A regression-based Monte Carlo method to solve backward stochastic differential equations*, Annals of Applied Probability, Vol.15(3), pp.2172-2002 (2005).
- [58] Benhamou, E., Gobet, E., Miri, M. : *Time dependent Heston model*, SIAM Journal on Financial Mathematics, Vol.1, pp.289-325, 2010.
- [59] Gulisashvili, A. : *A Asymptotic formulas with error estimates for call pricing functions and the implied volatility at extreme strikes*
- [60] Hamza, K., Klebaner, F.C. : *A family of non-Gaussian martingales with Gaussian marginals*, J. Appl. Math. and Stochastic Analysis, Article Id 92723, 2007.
- [61] Hagan, P.S., Kumar, D., Lesniewski, A.S., Woodward, D.E. : *Managing smile risk*, Willmott magazine, pp84-108, July 2012.
- [62] Hirsch, F., and Roynette, B. *A new proof of Kellerer's theorem*, ESAIM: Probability and Statistics 16:48-60, 2012.

- [63] Hirsch, F., Profeta, C., Roynette, B., Yor, M.: *Constructing self-similar martingales via two Skorohod embeddings*, Séminaire de Probabilités XLIII Lecture Notes in Mathematics 2011, pp 451-503.
- [64] Hirsch, F., Profeta, C., Roynette, B., Yor, M. : *Peacocks and Associated Martingales, with Explicit Constructions*, Springer (2011).
- [65] Hobson, D. : *Robust hedging of the lookback option*, Finance and Stochastics, 2, pp329-347, 1998.
- [66] Hobson, D. : *The Skorohod embedding problem and model-independent bounds for option prices*, In Paris-Princeton Lectures on Mathematical Finance 2010, volume 2003 of Lecture Notes in Math., pages 267318. Springer, Berlin, 2011.
- [67] Hobson, D. : *Fake Exponential Brownian Motion*, arXiv:1210.1391.
- [68] Hobson, D., Klimmek, M. : *Model independent hedging strategies for variance swaps*, Finance and Stochastics 16, 611-649, 2012.
- [69] Hobson, D., Klimmek, M. : *Robust price bounds for the forward starting straddle*, preprint, 2013.
- [70] Hirsch, F., Roynette, B. : *A new proof of Kellerer's theorem*, ESAIM: Probability and Statistics, Vol.16, Sept. 2012, pp 48-60.
- [71] Hirsch, F., Profeta, C., Roynette, B., Yor, M. : *Peacocks and Associated Martingales, with Explicit Constructions*, Springer (2011).
- [72] Hobson, D., Neuberger, A. : *Robust bounds for forward start options*, Mathematical Finance, Vol 22, Issue 1, pp 31-56 (Jan. 2012).
- [73] Hobson, D., Klimmek, M. : *Model independent hedging strategies for variance swaps*, to appear in Finance and Stochastics (2012).
- [74] Ikeda, N., Nagasawa, M., Watanabe, S. *Branching Markov Process I*, J. Math. Kyoto Univ., 8 (1968) 233-278.
- [75] Ikeda, N., Nagasawa, M., Watanabe, S. *Branching Markov Process II*, J. Math. Kyoto Univ., 8 (1968) 365-410.
- [76] Ikeda, N., Nagasawa, M., Watanabe, S. *Branching Markov Process III*, J. Math. Kyoto Univ., 9 (1969) 95-160.
- [77] Jacod, J. : *Calcul stochastique et problemes de martingales*, Spinrger, 1979.
- [78] Jacod, J., Shiryaev A.N. : *Limit theorems for stochastic processes*, Springer, 2003.
- [79] Jourdain, B., Sbai, M. : *Coupling Index and Stocks*, Quantitative Finance, October (2010).
- [80] Kahale, N. : *Model-independent lower bound on variance swaps*, preprint (2009).

- [81] Kellerer, H.G. : *Markov-Komposition und eine Anwendung auf Martingale*, Math. Ann., 198, 99122 (1972).
- [82] Karandikar, R.L. : *On pathwise stochastic integration*, Stoch. Proc. and their App. 57(1):11–18, 1995.
- [83] Kellerer, H.G. : *Markov-Komposition und eine anwendung auf martingale*, Math. Ann. 198:99–122, 1972.
- [84] McCann, R.J.: *A convexity principle for interacting gases*, advances in mathematics (128):153–179, 1997.
- [85] Korman, J. , McCann, R. : *Optimal transportation with capacity constraints*, To appear in Trans. Amer. Math. Soc, arXiv:1201.6404v2.
- [86] Lee, R. : *The Moment Formula for Implied Volatility at Extreme Strikes*, Mathematical Finance, Volume 14, Number 3, July (2004).
- [87] Le Jan, Y., Snitzman, A. : *Stochastic cascades and 3-dimensional Navier Stokes equations*, Prob. Theory Relat. Fields, 109, 343-366, (1997).
- [88] Li, P., Yau, S-T : *On the parabolic kernel of the Schrödinger operator*, Acta. Math., 156:153–201, 1986.
- [89] Lyons, T.J. : *Uncertain volatility and the risk-free synthesis of derivatives*, Applied Mathematical Finance, 2(2):117–133 (1995).
- [90] Madan, D.B., Yor, M. : *Making Markov martingales meet marginals: with explicit constructions*, Bernoulli, 8(4):509536, 2002.
- [91] McKean, H. P. : *Application of Brownian motion to the equation of Kolmogorov-Petrovskii-Piskunov*, Communications on Pure and Applied Mathematics, Vol. 28, Issue 3, pp 323-331, May 1975.
- [92] Méléard, S. : *Asymptotic behaviour of some interacting particle systems; McKean-Vlasov and Boltzmann models*, in Probabilistic Models for Non-linear Partial Differential Equations, Lecture Notes in Mathematics 1627, Springer-Verlag (1996).
- [93] Oblój, J. : *The Skorohod embedding problem and its offspring*, Probab. Surv., 1:321390, 2004.
- [94] Oblój, J. and Spoida, P. (2013), *An Iterated Azéma-Yor Type Embedding for Finitely Many Marginals*. Preprint, arXiv:1304.0368.
- [95] Oleszkiewicz, K. : *On fake Brownian motions*, Statistics and Probability Letters, 78, 12511254, 2008.
- [96] Osekowski, A. : *Sharp maximal inequalities for the martingale square bracket*, Stochastics An International Journal of Probability and Stochastic Processes, Volume 82, Issue 6, 2010.
- [97] Pardoux, E., Peng, S. : *Adapted Solution of a Backward Stochastic Differential Equation*, Systems Control Lett., 14, 55-61 (1990).

- [98] Pass, B.: *Uniqueness and Monge solutions in the multimarginal optimal transportation problem*, SIAM Journal of Mathematical Analysis, 43(6):2758-2775, 2011.
- [99] Pass, B. : *On a class of optimal transportation problems with infinitely many marginals*, preprint (Jun. 2012), arXiv:1206.5543v1.
- [100] Peccati, G., Yor, M. : *Burkholder's submartingales from a stochastic calculus perspective*, arXiv:0705.3633, preprint (2007).
- [101] Peskir, G., Shiryaev, A. N. : *Optimal Stopping and Free-Boundary Problems*, Lectures in Mathematics, Birkhäuser Basel (2006).
- [102] Rachev, S. T. and Ruschendorf, L. : *Mass Transportation Problems*. In Vol. 1: Theory. Vol. 2: Applications. Springer, Berlin, 1998.
- [103] Soner, M., Touzi, N., Zhang, J. : *Wellposedness of second order backward SDEs*, Probability Theory and Related Fields, 153, 149-190.
- [104] Skorohod, A.V. : *Branching diffusion processes*, Theory Probab. Appl., 9(3), 445-449, 1965.
- [105] Sznitman, A.S. : *Topics in propagation of chaos*, Ecole d'été de probabilités de Saint-Flour XIX - 1989, volume 1464 of Lect. Notes in Math. Springer-Verlag (1991).
- [106] Tan, X. and Touzi, N. : *Optimal Transportation under Controlled Stochastic Dynamics*, Annals of Probability, Vol. 41, No. 5, 3201-3240, 2013.
- [107] Villani, C. : *Topics in optimal transportation*, Graduate studies in Mathematics AMS, Vol 58.