Instituto Nacional de Matemática Pura e Aplicada

On a Parabolic Inverse Problem Arising in Quantitative Finance: Convex and Iterative Regularization

Author: **Adriano De Cezaro**

Advisor: **Prof. Dr. Jorge P. Zubelli**

Rio de Janeiro
April of 2010

To my wife Fabiana
and
my parents Luiz and Inês.

Certo dia no Pampa,
Ainda quando rapazito,
Tive que sair solito,
Para enfrentar o destino,
Pelo qual me apaixonei desde menino.

Pampa a fora,
Mundo a fora,
Nunca me cansei de buscar conhecimento,
E digo, mesmo aos que não tem talento,
É herança que não se estravia ao vento.

Sempre que volto para casa,
Estou de alma lavada,
Não me perdi em nenhuma encruzilhada,
Se hoje sou o que sou,
Muito devo a herança deixada.

Não quero passar em branco,
Para isso não me custa o esforço,
Por trabalho tenho gosto, por Matemática, admiração,
Tanto o que digo é verdade,
Que uma Matemática conquistou meu coração.

Agradecimentos

Muitas pessoas contribuiram para e elaboração deste trabalho. Dentre essas pessoas, quero dedicar especial agradecimento a minha esposa Fabiana, pelo incentivo de prosseguir sempre, pela dedicação, compreensão e amor durante essa jornada. Agradeço também a meus pais Luiz e Inês, por orientar o meu caminho e apoiar minhas decisões em todos os momentos. Aos demais familiares, os quais não posso citar todos uma vez que formam um conjunto quase não-enumerável.

Ao meu orientador Prof. Dr. Jorge P. Zubelli, meus sinceros agradecimentos por sua serenidade, paciência e competência demonstrados nesse período. Enfim, acredito ter aqui mais que um orientador, mas um grande amigo. Esta tese é um resultado de longas horas de trabalho conjunto que realizamos.

Também, deixo registrado os meus agradecimentos aos Professores Antônio Leitão (UFSC), Otmar Scherzer (Uni. Vienna) pela colaboração em trabalhos que de forma direta ou indireta fazem parte da construção desta tese.

A meus amigos e colegas, por fazerem jus a assim serem chamados. Com estes compartilhei bons minutos de conversa estimulante na sala do café e pelos corredores do IMPA. Não faço nomes para não esquecer de ninguém.

A banca examinadora pelos comentários e sugestões valiosas que aprimoraram a forma de apresentação desta tese.

Aos professores e funcionários do IMPA, meus agradecimentos por tudo o que me proporcionaram. Em especial, ao Ensino do IMPA pela forma prestativa e simpática com a qual sempre me receberam.

Ao CNPq pelo apoio financeiro entre 03/2006 a 06/2008.

Por fim, agradeço o apoio da direção, dos funcionários, dos amigos e colegas do IMEF - FURG.

Abstract

This thesis concerns the study of regularization strategies for the identification of the diffusion coefficient in a parabolic PDE arising in quantitative finance. Using the properties of the parameter-to-solution map, we address the use of convex and iterative regularization to the specific problem. We present a unified framework for the calibration of local volatility models which makes use of recent tools of convex regularization of ill-posed problems. The unique aspect of the present approach is to address the key issue of convergence and sensitivity analysis of the regularized solution, when the noise level of the observations goes to zero, in a general and rigorous way. In particular, we present convergence results that include convergence rates with respect to noise level in a fairly general context and goes beyond the classical quadratic regularization. Our approach directly relates to many of the different techniques that have been used in volatility surface estimation. In particular, it has direct connections with the Statistical concept of exponential families and entropy-based estimation.

Another aspect of this work concerns the iterative regularization of the local volatility calibration problem. The novelty of this approach is based on the existence of a local tangential cone condition. In this direction, we perform a rigorous convergence analysis of the nonlinear Landweber iteration and the adequacy of this to the local volatility calibration problem. Numerical tests confirm the iteration performance. Moreover, we rewrite the problem as a system of nonlinear ill-posed equations end explore iterative-Kaczmarz type regularization strategies.

Finally, we connect the convex regularization framework with the Financial concept of Convex Risk Measures.

Key words: local volatility surface identification, convex regularization, convergence rates, tangential cone condition, iterative regularization, convex risk measures.

Contents

Introduction

Results about existence, uniqueness and stability for the identification of coefficients in partial differential equations from indirect observations have many applications in problems arising in geophysics, medical diagnostics, quantitative finance, and other areas [13, 39, 58, 79, 87].

In general, these mathematical problems are nonlinear and ill-posed in the sense of Hadamard [48]. Moreover, in practical applications we only have access to data obtained by measurements. The sparsity of the measurements and the imperfection of models and measuring instruments imply that the data is inevitably corrupted by noise. Hence, these problems lead to challenging mathematical questions including issues of existence, uniqueness and stability of solutions. Open questions and interesting applications have attracted the attention of a considerable number of researchers from different areas of mathematics to the identification problems of coefficients in partial differential equations (PDEs).

Existence and uniqueness play an important role in identification problems. Indeed, they imply that we have enough data to determine the solution. In other words, the existence and uniqueness of the problem mean that the operator F that assigns the unknown coefficient to the solution of the PDE (in suitable spaces) is bijective. Moreover, in some cases, under natural constraints, uniqueness implies continuity of F^{-1} (stability). However, in many applications we have ill-posedness due to lack of continuity for F^{-1}. This creates serious numerical problems: If one wants, using a numerical method, to approximate the solution of a well-posed problem whose solution does not depend continuously on the data, then, in general, the numerical method becomes unstable. Differently, from the existence and uniqueness requirement, the stability cannot be recovered with “mathematical tricks”, such as imposing additional constraints.

An alternative for recovering stable solutions of ill-posed problems is to use the so-called *regularization methods*. These consist in the approximation of the ill-posed problem by a family of well-posed ones. Applying regularization methods to ill-posed problems constitutes a compromise between accuracy and stability. The modern regularization theory was developed throughout the 1950s and 1960s by Jhon and Tikhonov [81]. To this day, a vast number of publications have been written on regularization for inverse problems. See, for example, [10, 64] for linear inverse problems and [13, 39, 61, 79] for nonlinear inverse problems.

Coefficient identification in elliptic, parabolic and hyperbolic PDE models is a source of interesting inverse problems [41, 57, 58]. All these identification problems call for some type of regularization method. In particular, inverse problems in diffusion processes have enjoyed a good amount of attention from the mathematical community. See for instance [41]. Volatility identification (calibration) in the evolution of European option prices [15, 22, 33, 35, 55, 58] is a typical example of that kind of problems, and it is also of great relevance in mathematical finance. The first formulation of the direct problem in terms of PDEs was given by Black-Scholes [15] in 1973. In 1997, Scholes and Merton received the Nobel Prize in Economics for their contribution in this area.

This thesis is concerned with the theoretical aspects of the practical problem of calibrating (identifying) the volatility surface from the market-observed prices of European call options. This is a nonlinear ill-posed problem whose solution required regularization techniques. Here, we analyze Tikhonov with convex regularization and iterative regularization methods for the calibration problem in a European option price model.

We consider a complete financial market, where cash can be borrowed at a constant interest rate r, and a risky stock of value $S = S(t)$ that yields a continuously compounded dividend at a constant rate q, satisfying the diffusion price processes

$$dS(t) = S(t)\left(\nu(t, S(t))dt + \sigma(t, S(t))dW(t)\right), \quad t > 0, \quad S(0) = S_0 > 0, \tag{1}$$

where $W(t)$ denotes the standard Wiener process [63, 65]. The parameters ν and σ are called drift rate and underlying asset volatility, respectively.

A *European call option* with maturity date $T > 0$ and strike $K \geq 0$, on the underlying asset $S \geq 0$, consists of the right, but not the obligation, to buy, at a price K, a unit of S at time T. In the context of complete and arbitrage-free markets, the theoretical fair price, for the European call on S, has the probabilistic representation

$$U(0, S_0; T, K, r, q, \sigma^2) = \exp(-rT)\mathbb{E}_{\mathbb{Q}}^{0,S_0}(S(T) - K)^+, \tag{2}$$

where $\mathbb{E}_{\mathbb{Q}}^{0,S_0}$ is the expected value with respect to the *risk-neutral* probability measure $\mathbb{Q}$

given that, at $t = 0$, we have $S(0) = S_0$. Here, as usual, we define

$$(S - K)^+ := \max\{S - K, 0\} .$$

An interpretation of equation (2) would be: for each realization ω of the market, the payoff $(S(T,\omega) - K)^+$ should be brought to its present value $e^{-rT}(S(T,\omega) - K)^+$ by means of discounting by the interest rate r. Then, we average over all the possible realizations with respect to the risk-neutral measure $\mathbb{Q}$. The risk-neutral measure differs from the so-called subjective one in the sense that it is the one for which the discounted process S_t/e^{rt} is a martingale. See more details in [63, 65].

In this framework, the fair price for a European call option is given by the solution of the Black-Scholes partial differential equation [15]

$$U_t + \frac{1}{2}\sigma^2(t,S)S^2 U_{SS} + (r-q)SU_S - rU \ = 0\,, \quad t < T\,, \quad S \geq 0\,, \tag{3}$$

with final condition

$$U(t = T, S) = (S - K)^+ \, . \tag{4}$$

An important consequence of the Black-Sholes-Merton theory is that the drift rate ν in Equation (1) is not considered in (3). In fact, this lies at the core of the concept of the risk-neutral measure $\mathbb{Q}$.

In the instance where σ is a deterministic function of time or a constant, an explicit formula for the price U is well known as shown in the seminal paper [15] and also [52, 55]. In this context, a careful analysis of the theoretical volatility calibration problem was carried out in [36, 52, 55, 58].

On the other hand, it is also well know that the option price U depends on the maturity T and strike K too. This satisfies the, by now classical, *Dupire* forward equation [33]

$$-U_T + \frac{1}{2}\sigma^2(T,K)K^2 U_{KK} - (r-q)KU_K - qU = 0\,, \quad T > 0\,, \quad K \geq 0\,, \tag{5}$$

with the initial value

$$U(T = 0, K) = (S_0 - K)^+, \quad \text{for} \quad K > 0\,. \tag{6}$$

Dupire's equation is the starting point of our inverse problem analysis. As usual, the dividend and interest rates are known during the option life. If the volatility σ is known, this initial value problem is well-posed in the sense of Hadamard [48] and allows for the stable computation of the option price [58]. However, in practice, the volatility is not known explicitly, i.e., the crucial (unknown) parameter in the initial value problem determined by (5) and (6) is the volatility. Hence the importance of the inverse problem to be able to estimate or reconstruct the volatility function for observable market prices.

When performing the usual change of variables

$$K = S_0 e^y\,,\ \tau = T - t\,,\ b = q - r\,,\quad u(\tau, y) = e^{q\tau} U^{t,S}(T, K) \tag{7}$$

and

$$a(\tau, y) = \frac{1}{2}\sigma^2(T-\tau; S_0 e^y)\,, \tag{8}$$

in (5) and (6), this yields the Dupire equation with forward variables (τ, y)

$$-u_\tau + a(\tau, y)(u_{yy} - u_y) + b u_y = 0 \tag{9}$$

and initial condition

$$u(0, y) = S_0(1 - e^y)^+ \tag{10}$$

We shall make use of the following notation: Let $I \subset \mathbb{R}$ be an open (probably unbounded) interval and $1 \leq p \leq \infty$. We assume that $T > 0$ and use the notation $\Omega := (0, T) \times I$. Furthermore, as usual, $W_p^{1,2}(\Omega)$ denotes the space of functions $u(.,.)$ satisfying

$$||u||_{W_p^{1,2}(\Omega)} := ||u||_{L_p(\Omega)} + ||u_t||_{L_p(\Omega)} + ||u_y||_{L_p(\Omega)} + ||u_{yy}||_{L_p(\Omega)} < \infty\,.$$

For a fixed $\varepsilon > 0$ we denote by $\mathcal{H}(\Omega) := H^{1+\varepsilon}(\Omega)$.
We are concerned with the following

Definition 1. *Let $\overline{a} > \underline{a} > 0$ and $\underline{a} \leq a_0 \leq \overline{a}$ where $a_0 \in \mathcal{H}(\Omega)$ is known a priori. We define the admissible class of parameter by*

$$\mathcal{D}(F) := \{a \in a_0 + \mathcal{H}(\Omega) : \underline{a} \leq a \leq \overline{a}\}\,. \tag{11}$$

We remark that $\mathcal{D}(F)$ is a convex set.

For existence, uniqueness and some regularity estimates to the solution of (9) and (10) in $W_p^{1,2}(\Omega)$, for any $p \in [2, \tilde{p})$ for some $\tilde{p} < 3$, with $a \in \mathcal{D}(F)$, see, for example, [22, 35, 58] for a full demonstration or Chapter 1 for a sketch of the proof.

The Inverse Problem of European Option Prices

In the last decades, the option price inverse problem has attracted attracted a large amount of research and attention. See Section 1.6 and references cited therein as a starting point. The idea is, given a few assumptions about the underlying process, to attempt to derive the corresponding dynamics in the risk neutral measure from the option price observations. A correct derivation of the underlying process requires a complete knowledge of the volatility. The volatility is in fact the diffusion coefficient of the stochastic process (1). In a practical situation, it is not the local volatility $\sigma(T, K)$ that is known but the European option prices U themselves. Indeed, the local volatility is the unique quantity in (1), (9) and (10) that cannot be obtained from the market.

The nonlinear inverse problem of option pricing that we are concerned with is the identification (or calibration) of a local volatility surface $\sigma(T, K)$ by observations of the solutions

$$U^{T,K}(\cdot; r, q, \sigma) = U_*^{T,K}(T, S) \tag{12}$$

of (5) and (6) to match quoted market prices $U_*(T, K)$. Each observation is linked to the solution of (5) and (6) with different values of K and T.

The calibrated local volatility function is used by risk managers and traders to evaluate risk exposures, calculate the option risk sensitivities, such as delta and vega (the Greeks in financial literature) and the hedging instruments [45, 82]. If the model cannot price correctly European options, the hedging ratio will be difficult to calculate.

In practical situations, the price $U^{t,S}(T, K)$ is only known for a discrete set of maturities and strikes. Since we are interested in continuous observations of the price $U^{t,S}(T, K)$, this leads to an interpolation or an approximation that introduces noisy data u^δ, whose level δ is assumed to be known *a priori* and satisfies the inequality

$$||u^* - u^\delta||_{L^2(\Omega)} \leq \delta\,, \tag{13}$$

where u^* is the data associated to the actual value $a^* \in \mathcal{D}(F)$.

Therefore, given the framework developed earlier, the **inverse problem** that we are concerned with in this thesis is the calibration (identification) of the parameter

$$a^* = a^*(\tau, y) \in \mathcal{D}(F)\,, \tag{14}$$

solution of the operator equation

$$\begin{aligned} F\ :\mathcal{D}(F) \subset \mathcal{H}(\Omega) &\longrightarrow W_2^{1,2}(\Omega) \\ a \qquad &\longmapsto \quad F(a) = u(a) - u(a_0)\,, \end{aligned} \tag{15}$$

where $u(a)$ and $u(a_0)$ are solutions of (9) and (10) for $a, a_0 \in \mathcal{D}(F)$, given a set of data, probably corrupted by noise, that satisfies (13).

In this context, $a_0 \in \mathcal{D}(F)$ represents an *a-priori* in the Tikhonov regularization method in Chapter 2 and the starting point from iterative regularization methods studied in Chapter 3.

The thesis is organized as follows:

In the beginning of Chapter 1, we explain some properties of the direct problem of the Black-Scholes model to pricing European call options. In Subsection 1.2, we present some results of existence, uniqueness and some regularity estimates to the solution of the PDE (9) with the initial condition (10) and parameters in the admissible class $\mathcal{D}(F)$. These results are well known and can be found, for example, in [22, 35, 58]. Therefore, we only present a sketch of the proof. In Section 1.3, we prove properties of the parameter-to-solution map F that imply the ill-posedness of the inverse problem. In Theorem 1.3.1, we prove that the parameter-to-solution map F is continuous, compact and weakly closed. In Lemma 1.3.1, we prove the Fréchet differentiability of F.

In Section 1.4 we prove the new results about the parameter-to-solution map F. In Subsection 1.4.1, we characterize the kernel and the range of the Fréchet derivative of F. See Lemmas 1.4.1 and 1.4.2. These results are important for obtaining convergence rates for the Tikhonov regularization method proposed in Chapter 2 and Appendix A.

In Subsection 1.4.2 we present the main result of the Section 1.4. This is the verification of the **local tangential cone condition** to the parameter-to-solution map F defined in (15) for the calibration problem. This property is introduced in Theorem 1.4.2. The contribution of this result becomes apparent in the proof of convergence and stability of iterative regularization for the calibration problem in Chapter 3. The novel results of the chapter also appear in the articles [28, 29, 30] that have recently been submitted for publication.

At the end of this chapter, we briefly review some volatility calibration literature.

In Chapter 2 we propose the use of Tikhonov regularization by means of a convex regularizing functional as an extension to the quadratic regularization problem that has been used previously in the inverse problem literature [22, 23, 35]. Thus, we focus on inverse problem from the perspective of convex analysis methods and Bregman distances. On the theoretical side, our result yields better convergence rates and allows for convergence in spaces different from those in the quadratic regularization setting. For instance, it applies to Orlicz-Sobolev regularization [1]. In fact, in some cases, the convergence of certain convex regularization expressions implies convergence in the L^1-norm. Besides these results, our approach connects with central topics in different areas of current research. Such topics include *exponential families* of probability distributions, which is an important subject in Statistics.

In Section 2.1.1 we use recent convergence analysis results for nonlinear Tikhonov regularization [79] to prove existence and stability of the regularized solution of the calibration problem. In Subsections 2.1.3 and 2.1.4 we improve convergence rates in the literature to the specific inverse problem under consideration making use of results from Subsection 2.1.2. In inverse problems, to obtain convergence rate results, some *a-priori* information on the solution is needed. This *a-priori* information is called the *source condition.* From convex regularization, a natural source condition is

$$\xi^\dagger := F'(a^\dagger)^*\omega^\dagger \in \partial f(a^\dagger)\,, \tag{16}$$

where ∂f denotes the subdifferential of the convex functional f, as revised in the Appendix B. A heuristic financial interpretation of the source condition (16) is that we have a restriction that allows us to quantify the risk associated to a given volatility level. By this we mean that upon computing the corresponding Black-Scholes solution as a function of the volatility, we will be quantifying the level of risk one has in the space of random variables associated to such volatility. This is done with the help of the source condition (16) in Chapter 4. This way we construct a functional that, through the Fenchel duality, defines different convex risk measures. The availability of such risk measures permits quantifying the risk associated to random variables and portfolios of the underlying model.

The results in Subsection 2.1.2 help us prove the existence of an approximated source condition to the calibration problem.

In Section 2.2 we approach the regularization theory with convex penalization from a statistical point of view. In particular, we make use of the concept of *exponential families* of probability distributions to motivate Bregman distance as a regularization to the calibration problem. The connection between Bregman distances and exponential families is well-established in some contexts [2, 8], although in the present context our motivation in Section 2.2 is heuristic. This chapter is based on the accepted paper [29] and the submitted paper [28].

In Chapter 3, we analyze iterative regularization theory for the inverse problem (15), still focusing at local volatility surface calibration of the Black-Scholes model (9) and (10).

Iterative regularization of the local volatility calibration problem is novel to the best of our knowledge. This chapter's main contribution is the calibration of the local volatility surface by Landweber methods.

In Section 3.1, we review the convergence analysis of the Landweber iteration of nonlinear inverse problems. In Subsection 3.1.1, we verify that the assumptions of the classical Landweber iteration are satisfied by the parameter-to-solution map F as defined in (15). The implementation of the Landweber iteration in the $W_2^{1,2}$-inner product implies the evaluation of $F'(\cdot)^*$ with respect to this inner product for each iteration. The complexity of the inner product results in analytical and numerical difficulties evaluating $F'(\cdot)^*$. Hence, in Section 3.2, we recover the convergence analysis of the Landweber iteration using the discrepancy principle formulated in the $L^2(\Omega)$-norm. This simplifies the iteration implementation, for the calculation of $F'(\cdot)$ in the $W_2^{1,2}(\Omega)$-inner product is not an easy task. Some comments on the implementation of the nonlinear Landweber iteration are offered in Section 3.3. In this subsection we also present some numerical results that show the performance of the Landweber iteration to the calibration problem. In Section 3.4, we look at the calibration inverse problem as a system of nonlinear ill-posed equations. Kaczmarz type strategies [12, 11, 24, 50, 66] to solve nonlinear ill-posed problems are analyzed as possible iterative regularization methods to the calibration problem. This Chapter's results are collected from the working paper [30].

In Chapter 4, we demonstrate a connection between convex risk measures and the interpretation of source condition (16). The main point is a construction that allows us to allocate a convex risk measure to each convex regularization functional f involved in the source condition. This construction implies a financial interpretation of the source condition (16) as a restriction that allows us to quantify the risk associated to a given volatility level. This circle of ideas is novel and deserves careful further investigation of its financial and economical implications. In Section 4.1, we make use of the definition of the subdifferential and the source condition (16) to define a functional that, through the Fenchel conjugate, serves as our starting point to construct a convex risk measure in Section 4.2. In Subsection 4.2.1, we give an example of a specific convex risk measure associated with the Boltzmann-Shannon entropy. This chapter is based on the accepted paper [29] and the submitted paper [28].

In Chapter 5, we conclude with some remarks about the choice of the admissible parameters in relation to those that appear in the literature [22, 35]. We also make suggestions for future research related to the theory developed in this work.

In Appendix A, we provide a brief review of no-arbitrage results related to the volatility surface calibration proposed in this work. In Section A.2, we encourage the addition of a regularization term in the Tikhonov functional that will act as a penalization of arbitrage restriction on the shape of the option price surfaces. This restriction can be interpreted as a data preprocessing in the minimization. In Section A.3, using the penalization $f(a) = \|a - a_0\|^2_{\mathcal{H}(\Omega)}$ in the Tikhonov functional and properties of the calibration problem, we obtain better convergence rates of the regularized solution than those obtained in [22, 35], without using the general framework developed in Chapter 2.

In Appendix B, we present some definitions used in this work, in particular, definitions related to convex analysis used in Chapters 2 and 4 and concepts related to sufficient statistic and exponential families in Section 2.2.

Chapter 1

Direct *versus* Inverse Problem of Option Pricing

The new results contained in this chapter are shown in Section 1.4. In Subsection 1.4.1, we characterize the kernel and the range of the Fréchet derivative of F as $L^2(\Omega)$-subsets. See Lemmas 1.4.1 and 1.4.2. These results are important for obtaining convergence rates for the Tikhonov regularization method proposed in Chapter 2 and Appendix A.

The Theorem 1.4.2 in Subsection 1.4.2 is the main result of the Section 1.4. This theorem proves the **local tangential cone condition** to the parameter-to-solution map F defined in (15). The contribution of this result appears in the prove of convergence and stability of iterative regularization to the calibration problem in Chapter 3.

In financial markets a number of contracts are negotiated in such a way that their values are derived from other underlying assets or equities. Such derivative contracts play a fundamental role in risk management and corporate strategies. Their presence became so widespread that currently the volume of many derivative markets surpasses the value of the corresponding underlying markets.

The development of mathematical methods for pricing derivatives was a major reason for the expansion of derivative markets. Such theoretical achievement was recognized by the Nobel Prize in Economics award to R. Merton and M. Scholes. The corresponding methods involve the solution of the Black-Scholes partial differential equation, which in turn depends on the risk-free interest rate prevalent in the market, the dividend rate, and the volatility of the underlying asset. There are many models to describe the volatility. Among those, one that is very popular under practitioners is to assume that such volatilities are functions of the form $\sigma = \sigma(t, S)$, where t is the time and S is the asset price. It is usually referred to as Dupire's local volatility model [33] and $\sigma = \sigma(t, S)$ is called the *local volatility surface*.

Dupire's equation (5), with the initial value (6), is the starting point of our inverse problem analysis. As usual, the dividend and interest rates are known during the option life. If the volatility σ is known, this boundary value problem is well posed in the sense of Hadamard [48] and allows the stable computation of the option price. However, in practice, the volatility is not known explicitly, i.e., the crucial (unknown) parameter in the initial value problem determined by (5) and (6) is the volatility. Hence, the importance of the inverse problem, where one tries to estimate or reconstruct the volatility function for observable market prices.

1.1 The Direct Problem of European Option Pricing

For a constant volatility model, or a function of time alone, explicit formulas for the European option price U are well known [15, 36, 52, 55]. Indeed, it is the famous Black-Scholes formula [15]

$$U(S_0; T, K, r, q, \sigma^2) = S_0\mathcal{N}(d_1) - e^{-rt}K\mathcal{N}(d_2)\,, \tag{1.1}$$

with

$$d_1 = \frac{ln(s_0/K) + (r + \sigma^2/2)T}{\sigma\sqrt{T}}\,, \qquad d_2 = d_1 - \sigma\sqrt{T}\,,$$

and $\mathcal{N}(\cdot)$ denoting the distribution function of the standard normal distribution.

The price of a European call option as a function of volatility $U(\sigma)$ is strictly monotone (see [22]) and the unique value of σ corresponding to a given option price $U(K, T; S, t)$ is called the "Black-Scholes" *impled volatility*. In contrast to the constant parameter assumption of the Black-Scholes model, implied volatility shows a distinct dependence on the strike and maturity. This phenomena is referred as the *smile* effect [33, 34]. A possibility to explain this phenomena is by using the deterministic function $\sigma(K, T)$ in Dupire's option price PDE.

Below we do a brief review about existence and uniqueness of solutions of parabolic equations like (9) and (10) in Sobolev spaces. The presented results are collected from [22, 35, 58].

1.2 Existence and Uniqueness of the Parabolic Problem

Initially, our focus is to show existence and uniqueness of a solution u to the parabolic PDE (9) and (10). This is done following some ideas in [22, 35, 58, 70].

Remark 1.2.1. *In [22, 35], they consider the admissible class of parameters as a subset of $H^1(\Omega)$. We return to this question in Chapter 5.*

Our particular focus here concerns on existence and uniqueness in the space $W_2^{1,2}(\Omega)$. But it is well known that the results which follow are true in $W_p^{1,2}(\Omega)$ for some $p \in [2, 3)$. See, for example, [22, 35]. The results are based on extending, by density, the well known results in parabolic PDEs with Hölder continuous coefficients [58, 70].

Proposition 1.2.1. *[70, Theorem IV 9.2] Let a be Hölder continuous with $\underline{a} \le a \le \overline{a}$, $b \in L^\infty(\Omega)$ and $f \in L^p(\Omega) \cap L^2(\Omega)$. Then*

$$- v_\tau + a v_{yy} + b v_y = f \tag{1.2}$$

$$v(0, y) = 0 \tag{1.3}$$

has a unique solution $v \in W_2^{1,2}(\Omega)$. Moreover,

$$\|v\|_{W_2^{1,2}(\Omega)} \le C \|f\|_{L^2(\Omega)}, \qquad \textit{were} \quad C = C(\|a\|_{L^\infty(\Omega)}, \|b\|_{L^\infty(\Omega)}). \tag{1.4}$$

Using density embedding theorems for Sobolev spaces it follows:

Proposition 1.2.2. *[35, Proposition A1] Let $a \in \mathcal{D}(F)$, $b \in L^\infty(\Omega)$ and $f \in L^p(\Omega) \cap L^2(\Omega)$. Then, the initial value problem* (1.2) *and* (1.3) *has a unique solution $v \in W_2^{1,2}(\Omega)$. Moreover, it satisfies the regularity estimate* (1.4).

Here we present only the proof of existence of solutions in $W_2^{1,2}(\Omega)$. For a complete proof, see [35, Proposition A1].

Given a Hölder coefficient a_n, Proposition 1.2.1 implies that there exists a unique solution $v^n \in W_2^{1,2}(\Omega)$ of

$$\begin{aligned} -v^n_\tau + a_n v^n_{yy} + b v^n_y &= f \\ v(0, y) &= 0 \end{aligned} \tag{1.5}$$

that satisfies the estimate

$$\|v^n\|_{W_2^{1,2}(\Omega)} \le C\|f\|_{L^2(\Omega)} . \tag{1.6}$$

Let $w^n := v^n_y$. By linearity of equation (1.5) its satisfies

$$\begin{aligned} -w^n_\tau + (a_n w^n_y)_y + (bw^n)_y &= f_y \\ w^n(0,y) &= 0 \end{aligned}$$

with the estimate

$$\|w^n\|^2_{W_2^{0,1}(\Omega)} \le c_1 \int_0^T \|f_y(t)\|^2_{W_2^{-1}(\mathbb{R})} dt \le c_2 \|f\|_{L^2(\Omega)} , \tag{1.7}$$

where c_1 and c_2 depend on the limitation of the coefficients.

From the definition of $W_2^{1,2}(\Omega)$-norm

$$\|v^n\|_{L^2(\Omega)} \le C\|f\|_{L^2(\Omega)} .$$

Thus (1.6) holds uniformly for smooth coefficients $a_n \in \mathcal{D}(F)$.

Let $a \in \mathcal{D}(F)$. By density of Hölder spaces on $\mathcal{H}(\Omega)$ [1], there exists a sequence of Hölder coefficients a_n such that $a_n \to a \in \mathcal{D}(F)$ (in the $\mathcal{H}$-norm). Weak compactness of Hilbert spaces [84, 86] implies that there exists a subsequence v^{n_k} of v^n such that $v^{n_k} \rightharpoonup \hat{v} \in W_2^{1,2}(\Omega)$. Take $\psi \in C_0^\infty(\Omega)$. Hence

$$\int_\Omega (-\hat{v}_\tau + a\hat{v}_{yy} + b\hat{v}_y)\,\psi d(y,\tau) = \lim_{k\to\infty} \int_\Omega \left(-v^{n_k}_\tau + av^{n_k}_{yy} + bv^{n_k}_y\right)\psi d(y,\tau) = \int_\Omega f\psi d(y,\tau) .$$

By the weakly lower semi-continuity of $W_2^{1,2}(\Omega)$-norm $\hat{v}$ satisfies (1.7). Hence, $v := \hat{v}$ is the unique weak solution of (1.2) with initial condition (1.3). □

In the following we give a representation of the solution of (9) and (10).

Corollary 1.2.1. *[35, Corollary A1] Let $a \in \mathcal{D}(F)$, $b \in L^\infty(\Omega)$. Then, there exists a unique solution $u \in W^{1,2}_{2,loc}(\Omega)$ to the problem* (9) *and* (10) *that satisfies*

$$\|u\| \le S_0 \quad and \quad \|u_y\|_{W_2^{0,1}(\Omega)} \le C , \tag{1.8}$$

were $c = c(\underline{a}, \|a\|_{L^\infty(\Omega)}, \|a\|_{L^\infty(\Omega)})$ and $C = C(\underline{a}, \|a\|_{L^\infty(\Omega)}, \|a\|_{L^\infty(\Omega)})$.

Sketch of the proof. Define

$$\bar{u}(\tau, y) := \int_{-\infty}^{0} \frac{1}{\sqrt{4\pi\tau}} e^{-\frac{(y-\theta)^2}{4\tau}} S_0(1-e^\theta)^+ d\theta \quad \in W^{1,2}_{2,loc}(\Omega)\,.$$

Take $w = u - \bar{u}$. By linearity, w solves

$$-w_\tau + a(w_{yy} - w_y) + bw_y = \bar{u}_\tau - a(\bar{u}_{yy} - \bar{u}_y) + b\bar{u}_y\,,$$

with homogeneous boundary conditions. Proposition 1.2.2 concludes the assertion. For more details see [35, Corollary A1]. □

Proposition 1.2.2 and Corollary 1.2.1 hold true in $W^{1,2}_p(\Omega)$ for $p \in [2, \tilde{p})$ with some $\tilde{p} > 2$. Moreover, with similar arguments the above results remain true for $a \in \mathcal{D}(F) \subset H^1(\Omega)$. For more details see [22, 35].

The fact that the parameter-to-solution map $F : a \longrightarrow u(a) - u(a_0)$ is well defined follows from linearity of (9) and Corollary 1.2.1.

The following lemma will be useful in the next section.

Lemma 1.2.1. *[22, Proposition 4.4 (1)] Assume that* $p_1 \in]2, \tilde{p}[$. *Let* $u \in W^{1,2}_{p_1}(\Omega)$ *be a solution of* (9) *and* (10). *Then, there exists a constant* C *that depends only on the boundedness of the coefficients and* p_1 *such that*

$$\|u_{yy} - u_y\|_{L_{p_1}(\Omega)} \le C\,. \tag{1.9}$$

Sketch of the proof. A complete proof can be see in [22, Proposition 4.4 (1)]. □

1.3 Known Properties of the Forward Operator

In this section, we summarize some properties of the parameter-to-solution map F defined by (15). The results in this section are collected from [22, 35]. In [22], the results are based on probabilistic arguments. After that, in [35], similar results are obtained using a PDE approach. In particular, Theorem 1.3.1 and Lemma 1.3.1 are particular cases of [35, Theorem 2.1] and [35, Proposition 4.1]. Analogous results appeared also in [22, Proposition 5.1].

In this thesis we use the following definition of compactness:

Definition 1.3.1. $F : \mathcal{D}(F) \subset U \to V$ *is compact if for every bounded sequence* (x_k) *in*

$\mathcal{D}(F)$, $(F(x_k))$ has a convergent subsequence.

In particular, the composition of a compact linear operator and a sequentially continuous (not necessarily linear) operator is compact.

The next result implies in compactness and weak closedness of F.

Theorem 1.3.1. *[35, Theorem 2.1] The operator $F : \mathcal{D}(F) \subset \mathcal{H}(\Omega) \to W_2^{1,2}(\Omega)$ is compact. Moreover, F is weakly (sequentially) continuous and thus weakly closed.*

Sketch of the proof. We present only a simple version of the proof. For a complete version, see [35, Theorem 2.1]. Let $u^n = u(a_n)$ and $u = u(a)$ with $a_n, a \in \mathcal{D}(F)$ satisfying $a_n \rightharpoonup a$ (in $\mathcal{H}$-topology) and $u^n \rightharpoonup u$ (in $W_2^{1,2}$-topology). Let $\Omega_c \subset \Omega$ compact. By the Sobolev compact embedding Theorem [1, Theorem 4.12, case B, pg 85]

$$a_n \to a \quad \in L^q(\Omega_c) \qquad 1 \leq q < \infty\,. \tag{1.10}$$

By linearity $w^n := u^n - u$ satisfies

$$-w^n_\tau + a_n(w^n_{yy} - w^n_y) + bw^n_y = -(a - a_n)(u_{yy} - u_y)\,, \tag{1.11}$$

with homogeneous boundary conditions.

Take $\Omega_c = (0,T) \times (-M, M)$ and $\Omega_c^c = \Omega - \Omega_c$. So, equation (1.11) can be rewriting as

$$-w^n_\tau + a_n(w^n_{yy} - w^n_y) + bw^n_y = -(a - a_n)(u_{yy} - u_y) \cdot \chi_{\Omega_c} - (a - a_n)(u_{yy} - u_y) \cdot \chi_{\Omega_c^c}\,.$$

Proposition 1.2.2 implies the estimate

$$\|w^n\|_{W_2^{1,2}(\Omega)} \leq c\left(\|a_n - a\|_{L^q(\Omega_c)}\|u_y\|_{W_p^{1,1}(\Omega_c)} + \|a_n - a\|_{L^q(\Omega_c^c)}\|u_y\|_{W_p^{1,1}(\Omega_c^c)}\right) \tag{1.12}$$

for $\frac{1}{2} + \frac{1}{q} = \frac{1}{p}$.

Equations (1.8) and (1.10) imply that the first term on the right hand side of (1.12) goes to zero with $n \to \infty$. Moreover $\|u_y\|_{W_p^{1,1}(\Omega_c^c)} \to 0$ as $M \to \infty$. Consequently, the second term in the estimate (1.12) goes to zero as a suitable choice of M. This concludes the assertion. □

The next lemma guarantees the existence of the one-sided directional derivative of F for $a \in \mathcal{D}(F)$ in the direction h such that $a + h \in \mathcal{D}(F)$. For a definition of an one-side directional derivatives see [84, 21].

Lemma 1.3.1. *[35, Proposition 4.1] Consider the operator F as in Theorem 1.3.1. Then, F admits a one side derivative at $a \in \mathcal{D}(F)$ in the direction h such that $a+h \in \mathcal{D}(F)$. The derivative $F'(a)$ satisfies $\|F'(a)h\|_{W_p^{1,2}(\Omega)} \leq c\|h\|_{\mathcal{H}(\Omega)}$. In other words, $F'(a)$ is extendable to a bounded linear operator on $\mathcal{H}(\Omega)$. Moreover, $F'(a)$ satisfies the Lipschitz condition*

$$\|F'(a) - F'(a+h)\|_{\mathcal{L}(\mathcal{H}(\Omega), W_p^{1,2}(\Omega))} \leq \gamma\|h\|_{\mathcal{H}(\Omega)} \ , \tag{1.13}$$

for all a and h such that $a,\ a+h\ \in \mathcal{D}(F)$.

Sketch of the proof. The proof follows similar arguments of the full proves in [35, Proposition 4.1]. Let $a \in \mathcal{D}(F)$ and the direction $h \in \mathcal{H}(\Omega)$ be such that $a+h \in \mathcal{D}(F)$. For simplicity of exposition, let us assume that $b = 0$ in (9) and (10). By the linearity of equation (9) the directional derivative $u' \cdot h$ in the direction h satisfies

$$-(u' \cdot h)_\tau + a((u' \cdot h)_{yy} - (u' \cdot h)_y) = -h(u_{yy} - u_y) \tag{1.14}$$

with homogeneous initial conditions. From Proposition 1.2.2 there exists a single solution $u' \cdot h \in W_p^{1,2}(\Omega)$ of (1.14), for $2 \leq p < \bar{p}$.

Using regularity estimates to the parabolic problem (see for example [70]) we have

$$\begin{aligned}\|u' \cdot h\|_{W_p^{1,2}(\Omega)} &\leq c\|h(u_{yy} - u_y)\|_{L_p(\Omega)} \\ &\leq c\|h\|_{L_{p_2}(\Omega)}\|(u_{yy} - u_y)\|_{L_{p_1}(\Omega)} \ ,\end{aligned} \tag{1.15}$$

where $p_1 \in (p, \bar{p})$ and p_2 satisfies $1/p = 1/p_1 + 1/p_2$. Note that, $p_2 = \frac{p_1 p}{p_1 - p}$. Lemma 1.2.1 implies that $\|u_{yy} - u_y\|_{L_{p_1}(\Omega)} \leq C$ for all $a \in \mathcal{D}(F)$.Moreover, from the Sobolev Embedding Theorem [1, Theorem 4.12, case B, pg 85] it follows that there exists a constant $c > 0$ such that $\|h\|_{L_{p_2}(\Omega)} \leq c\|h\|_{\mathcal{H}(\Omega)}$, for all $h \in \mathcal{H}(\Omega)$. Now, equation (1.15) implies that

$$\|u'(a) \cdot h\|_{W_p^{1,2}(\Omega)} \leq C\|h\|_{\mathcal{H}(\Omega)} \, . \tag{1.16}$$

Thus, the derivative $u'(a) = F'(a)$ can to be extended as a bounded linear operator to $\mathcal{H}(\Omega)$. The next step is to obtain the Lipschitz condition (1.13). To do this, denote by $\tilde{u}(\tilde{a})$ the solution of (9) and (10) with a replaced by $\tilde{a} = a + h$ with $h \in \mathcal{H}(\Omega)$. Setting $v := (F'(\tilde{a}) - F'(a)) \cdot q = (\tilde{u'} - u') \cdot q$ with $q \in \mathcal{H}(\Omega)$. Then, from linearity of (9), v is a solution of

$$(v)_\tau + a((v)_y - (v)_{yy}) = q((\tilde{u} - u)_{yy} - (\tilde{u} - u)_y) + (\tilde{a} - a)((\tilde{u'} \cdot q)_{yy} - (\tilde{u'} \cdot q)_y) \ . \quad (1.17)$$

Using an estimate analogous to (1.16) we find

$$\begin{aligned} \|v\|_{W_p^{1,2}(\Omega)} &\le (\tilde{c}\|q\|_{\mathcal{H}(\Omega)}\|\tilde{u} - u\|_{W_{\bar{p}}^{1,2}(\Omega)} + \bar{c}\|\tilde{a} - a\|_{\mathcal{H}(\Omega)}\|\tilde{u'} \cdot q\|_{W_{\bar{p}}^{1,2}(\Omega)}) \\ &\le C\|q\|_{\mathcal{H}(\Omega)}\|\tilde{a} - a\|_{\mathcal{H}(\Omega)} \ . \end{aligned}$$

Taking the sup over all $q \in \mathcal{H}(\Omega)$ satisfying $\|q\|_{\mathcal{H}(\Omega)} \le 1$, on both sides of the above inequalities we have the Lipschitz condition (1.13). □

As observed in [35, Remark 4.1], $\mathcal{D}(F)$ has no interior points when equipped with the $H^1(\Omega)$ norm. Because of that, $F'(a)$ is not necessarily differentiable in every direction $h \in H^1(\Omega)$. In other words, $F'(a)$ is not Gateaux differentiable. This will not affect the convergence analysis that follows. In fact, for such analysis we only need that the operator F attains a one-sided directional derivative at $a^\dagger$ in the directions $a - a^\dagger$ for all $a \in \mathcal{D}(F)$. The sufficient condition for this to happen is $\mathcal{D}(F)$ to be starlike with respect to $a^\dagger$. That is, for every $a \in \mathcal{D}(F)$ there exists $t_0 > 0$ such that

$$a^\dagger + t(a - a^\dagger) = ta + (1 - t)a^\dagger \in \mathcal{D}(F) \qquad \forall 0 \le t \le t_0 \ .$$

As $\mathcal{D}(F)$ is convex, the requirement above follows. Moreover, the bounded linear operator $F'(a^\dagger)$ has properties that mimic the Gâteaux derivative.

In particular, there exists an adjoint operator

$$F'(a^\dagger)^* : V \longrightarrow U$$

defined by

$$\langle F'(a^\dagger)^* v, a\rangle_{L^2} = \langle v, F'(a^\dagger)a\rangle_{\mathcal{H}(\Omega)} \ , \quad a \in \mathcal{H}(\Omega) \ , \quad v \in V \ .$$

We emphasize that Theorem 1.3.1 holds true if we restrict our attention to

$$\mathcal{D}(F) := \{a \in a_0 + H^{1+\varepsilon}(\Omega) : \underline{a} \le a \le \overline{a}\}. \quad (1.18)$$

and a convex, weakly lower semi-continuous functional f on $\mathcal{H}(\Omega)$ with $\mathcal{D}(F) \subseteq \mathcal{D}(f)$. Moreover, for $\varepsilon > 0$, by the Sobolev embedding theorem, each function of $\mathcal{D}(F) \subset \mathcal{H}(\Omega)$

is an interior point, for which Fréchet-differentiability holds, as Lemma 1.3.1 shows.

1.4 New Results and Properties of the Forward Operator

In this section we show the new results obtained to the parameter-to-solution map F as defined in (15). In particular, the results in this section represent our main contribution in the development of convergence rates to Tikhonov regularization in Chapter 2 and in the analysis of stability and convergence for iterative calibration of the volatility surface in Chapter 3.

We start the section with the proof of the properties of F obtained before in Theorem 1.3.1 in the context of the $L^2(\Omega)$-norm.

Theorem 1.4.1. *The operator $F : \mathcal{D}(F) \subset \mathcal{H}(\Omega) \to L^2(\Omega)$ in Equation 15 is continuous and compact. Moreover, F is sequentially weakly continuous and weakly closed.*

Proof. The proof follows from Theorem 1.3.1, where it is proven that $F : \mathcal{D}(F) \subset \mathcal{H}(\Omega) \to W^{1,2}_p(\Omega)$ satisfies the property for all $2 \leq p < \bar{p}$ with an appropriate $\bar{p} > 2$. The result then follows by using that the embedding from $W^{1,2}_p(\Omega)$ into $L_2(\Omega)$ is linear and bounded. □

1.4.1 An $L^2(\Omega)$ Characterization of $\mathcal{R}(F'(a))$ and $\mathcal{N}(F'(a)^*)$

Below, we characterize the image of the operators $F'(\cdot)$ and $F'(\cdot)^*$ as $L^2(\Omega)$ subsets. The contribution of these results appears in the proof of convergence rates of Tikhonov regularization at Chapters 2 and Appendix A, respectively.

Lemma 1.4.1. *Let $a \in \mathcal{D}(F)$. Then, the Fréchet derivative of the operator F is injective and compact.*

Proof. Let $h \in \mathcal{N}(F'(a)) \subset \mathcal{H}(\Omega)$. Because of equation (1.14) we have that

$$h \cdot (u_{yy} - u_y) = 0\,, \tag{1.19}$$

where u is the solution of (9) and (10). However, $G(\tau, y) = (u_{yy} - u_y)$ is the distributional solution of the initial value problem

$$\begin{aligned} \partial_\tau G(\tau, y) &= \frac{1}{2}(\partial_y^2 - \partial_y)(a(t,y)G(\tau,y)) + b\,G(\tau,y) \\ G(0,y) &= \delta(y)\,, \end{aligned} \tag{1.20}$$

where $\delta(y)$ is the Dirac's delta. In others words, $G(\tau, y)$ is the Green's function of the Cauchy problem (1.20). Hence, $G(\tau, y) > 0$ for every y and $\tau > 0$ (See [58, Theorem 9.3.1 pg 271]). Thus, it follows from (1.19) that $h = 0$ a.e. The compactness of $F'(a)$ follows from the compactness of F. □

Lemma 1.4.2. *The operator $F'(a^\dagger)^*$ has a trivial kernel.*

Proof. As before, we take $b = 0$ for simplicity. Denote by

$$\mathcal{L}_u := -\partial_\tau + a(\partial_{yy} - \partial_y)$$

and by $G_{u_{yy}-u_y}$, the parabolic partial differential operator on the left hand side of Equation (1.14) with homogeneous boundary condition and the multiplication operator by the function $u_{yy} - u_y$, respectively. Hence, the solution of (1.14) has a functional form $u'(a) := F'(a) = \big(\mathcal{L}_u\big)^{-1} G_{u_{yy}-u_y}$, where by $\big(\mathcal{L}_u\big)^{-1}$ we mean the left-inverse of the operator $\mathcal{L}_u$ with vanishing boundary and initial conditions.

From the definition of $F'(a^\dagger)^* : V \longrightarrow \mathcal{H}(\Omega)$, we have

$$\langle F'(a^\dagger)h, z\rangle_V = \langle h, \varphi\rangle_{\mathcal{H}}(\Omega)\,, \qquad \forall h \in \mathcal{H}(\Omega)\,, \forall z \in V$$

and $F'(a^\dagger)^* z = \varphi$. Now, let $z \in \mathcal{N}(F'(a^\dagger)^*)$. Then,

$$\begin{aligned} 0 = \langle F'(a^\dagger)h, z\rangle_V &= \langle \big(\mathcal{L}_u\big)^{-1} G_{u_{yy}-u_y} h, z\rangle_V = \langle G_{u_{yy}-u_y} h, \big(\big(\mathcal{L}_u\big)^{-1}\big)^* z\rangle_V \\ &= \langle G_{u_{yy}-u_y} h, g\rangle_V = \int_\Omega (u_{yy} - u_y) h\, g\, d\tau\, dy \qquad \forall h \in \mathcal{H}(\Omega)\,. \end{aligned}$$

where g is a solution of the adjoint equation

$$g_\tau + (a^\dagger g)_{yy} + (a^\dagger g)_y = z\,,$$

with homogeneous final and boundary conditions. Since $z \in V = L^2(\Omega)$, $g \in \mathcal{H}(\Omega)$. See

[70]. In particular

$$\int_\Omega (u_{yy} - u_y) h \, g d\tau \, dy = 0 \,,$$

holds true for $h = g$. Since $G_{u_{yy}-u_y} > 0$ (see the end of the proof of Lemma 1.4.1) it follows that $g = 0$. Consequently, $z = 0$ and $\mathcal{N}(F'(a^\dagger)^*) = \{0\}$. □

Remark 1.4.1. *The range of $F'(a^\dagger)^*$ is dense in $\mathcal{H}(\Omega)$. Indeed, $\mathcal{H}(\Omega) = \overline{\mathcal{R}(F'(a^\dagger)^*)}^{H^{1+\varepsilon}} \oplus \mathcal{N}(F(a^\dagger))$ and the claim follows from Lemma 1.4.1.*

1.4.2 The Local Tangential Cone Condition for the Calibration Problem

In this subsection we prove the main contribution of this thesis in order to characterize properties of the parameter-to-solution map F. The importance of the result presented below is that, given the local tangential cone condition, we are able prove stability and convergence of iterative regularization for the calibration problem in Chapter 3. This is novel to the best of our knowledge.

We remark that, a consequence of Lemma 1.4.1 is that F cannot be constant along any affine subspace through a and parallel to $\mathcal{N}(F'(a))$. This means that the *tangential cone condition* (see (1.21) below) is not still a severe requirement. See the commentaries in [39, Chapter 11]. The next result is one of our contributions in iterative methods to the volatility calibration inverse problem. It is one of the main results of this section.

Theorem 1.4.2. *The parameter-to-solution map $F : \mathcal{D}(F) \subset \mathcal{H}(\Omega) \to W_2^{1,2}(\Omega)$ satisfies the local tangential cone condition*

$$\|F(a) - F(\tilde{a}) - F'(\tilde{a})(a - \tilde{a})\|_{W_2^{1,2}(\Omega)} \leq \eta \|F(a) - F(\tilde{a})\|_{W_2^{1,2}(\Omega)} \,, \qquad \eta < \frac{1}{2} \,, \tag{1.21}$$

for all $a, \tilde{a}$ in a ball $B_\rho(a^*) \subset \mathcal{D}(F)$ with some $\rho > 0$.

Proof. For easiness we consider $b = 0$ in (9). Denote by $u = u(a) = F(a)$ and $\tilde{u} = u(\tilde{a}) = F(\tilde{a})$ the solutions of (9) with $a, \tilde{a} \in \mathcal{D}(F)$. Moreover, let $\tilde{u}'h$ be the directional derivative of $F(\tilde{a})$ in the direction $h = a - \tilde{a}$ (that exist from Lemma 1.3.1). It satisfies the partial differential equations (9) and (1.14), for $a, \tilde{a} \in \mathcal{D}(F)$, respectively. Subtracting the last two equations from the first, it satisfies the PDE

$$-u_\tau + a(u_{yy} - u_y) + \tilde{u}_\tau - \tilde{a}(\tilde{u}_{yy} - \tilde{u}_y) + (\tilde{u}' \cdot h)_\tau - \tilde{a}((\tilde{u}' \cdot h)_{yy} - (\tilde{u}' \cdot h)_y) = h(\tilde{u}_{yy} - \tilde{u}_y) \,.$$

Rewritten, we have

$$\begin{aligned}-(u_\tau - \tilde{u}_\tau - (\tilde{u}' \cdot h)_\tau) + \frac{(a+\tilde{a})}{2}(u_{yy} - u_y) + \frac{(a-\tilde{a})}{2}(u_{yy} - u_y) - \frac{(a+\tilde{a})}{2}(\tilde{u}_{yy} - \tilde{u}_y) \\ + \frac{(a-\tilde{a})}{2}(\tilde{u}_{yy} - \tilde{u}_y) - \frac{(a+\tilde{a})}{2}((\tilde{u}' \cdot h)_{yy} - (\tilde{u}' \cdot h)_y) \\ + \frac{(a-\tilde{a})}{2}((\tilde{u}' \cdot h)_{yy} - (\tilde{u}' \cdot h)_y) = h(\tilde{u}_{yy} - \tilde{u}_y)\end{aligned}$$

or,

$$\begin{aligned}-(u_\tau - \tilde{u}_\tau - (\tilde{u}' \cdot h)_\tau) + \frac{(a+\tilde{a})}{2}\left[(u_{yy} - u_y) - (\tilde{u}_{yy} - \tilde{u}_y) - ((\tilde{u}' \cdot h)_{yy} - (\tilde{u}' \cdot h)_y)\right] \\ = -\frac{(a-\tilde{a})}{2}(u_{yy} - u_y) - \frac{(a-\tilde{a})}{2}((\tilde{u}' \cdot h)_{yy} - (\tilde{u}' \cdot h)_y) \\ - \frac{(a-\tilde{a})}{2}(\tilde{u}_{yy} - \tilde{u}_y) + h(\tilde{u}_{yy} - \tilde{u}_y)\,.\end{aligned}$$

Now consider the direction $h = a - \tilde{a}$. By linearity $w = u - \tilde{u} - \tilde{u}'(a - \tilde{a})$ satisfies the PDE

$$\begin{aligned}-w_\tau + \frac{(a+\tilde{a})}{2}(w_{yy} - w_y) = -\frac{(a-\tilde{a})}{2}\left[(\tilde{u}'h)_{yy} - (\tilde{u}'h)_y\right] \\ - \frac{(a-\tilde{a})}{2}(u_{yy} - u_y) + \frac{(a-\tilde{a})}{2}(\tilde{u}_{yy} - \tilde{u}_y)\end{aligned} \tag{1.22}$$

with homogeneous initial and boundary conditions.

From (1.17)

$$\begin{aligned}-\frac{(a-\tilde{a})}{2}\left[(\tilde{u}'h)_{yy} - (\tilde{u}'h)_y\right] = -\frac{1}{2}\left[-(\tilde{u}'h - u'h)_\tau + a((\tilde{u}'h - u'h)_{yy} - (\tilde{u}'h - u'h)_y)\right] \\ - \frac{h}{2}\left[(\tilde{u} - u)_{yy} - (\tilde{u} - u)_y\right]\,.\end{aligned}$$

Substituting the above equation into (1.22) we have

$$\begin{aligned}-w_\tau + \frac{a+\tilde{a}}{2}(w_{yy} - w_y) = -\frac{1}{2}\left(-[(\tilde{u} - u)'h]_\tau + a\left[((\tilde{u} - u)'h)_{yy} - ((\tilde{u} - u)'h)_y\right]\right) \\ - 2\frac{a-\tilde{a}}{2}\left((\tilde{u} - u)_{yy} - ((\tilde{u} - u))_y\right)\end{aligned}$$

Note that the right hand side of the above equation is in $L^2(\Omega)$. Hence, using a regularity estimate similar to (1.16) and the continuous Sobolev embedding Theorem [1]

$$
\begin{aligned}
\|w\|_{W_2^{1,2}(\Omega)} &\le c\left(\frac{1}{2}\|(\tilde{u}-u)'h\|_{W_2^{1,2}(\Omega)} + \|a-\tilde{a}\|_{L^2(\Omega)}\|\tilde{u}-u\|_{L^2(\Omega)}\right) \\
&\le C\left(\|(\tilde{u}-u)'h\|_{W_2^{1,2}(\Omega)} + \|a-\tilde{a}\|_{\mathcal{H}(\Omega)}\|\tilde{u}-u\|_{W_2^{1,2}(\Omega)}\right) \\
&\le C_1\|a-\tilde{a}\|_{\mathcal{H}(\Omega)}\|\tilde{u}-u\|_{W_2^{1,2}(\Omega)}\,,
\end{aligned}
\tag{1.23}
$$

where $C_1 = C_1(\underline{a}, \|a\|_{L^\infty(\Omega)}, \overline{a})$.

Let $\rho > 0$ be small enough such that $\eta := C_1\|a-\tilde{a}\|_{\mathcal{H}(\Omega)} < \frac{1}{2}$ for all $a, \tilde{a} \in B_\rho(a^*) \subset \mathcal{D}(F)$. Then, from (1.23) the tangential cone condition (1.21) is satisfied. □

The local tangential cone condition (1.21) and the scaled property (1.16) for all $a\,,\, \tilde{a} \in B_{2\rho}(a^*) \subset \mathcal{D}(F)$ are strong enough to ensure the local convergence of the iterative regularization method in Chapter 3 to a solution of (15), if (15) is solvable in $B_\rho(a^*)$. They also guarantee that the iterative methods are well defined in a suitable neighborhood of a^*.

1.5 Ill-Posedness of the Inverse Problem

For a given (incomplete) set of option prices, it is important to determine what information about the structure of the local volatility can be recovered in a unique and stable way. That is, how much information is needed for the calibrated local volatility surface be well-posed in Hadamard's sense [48]. We emphasize that stability plays a crucial role for a robust calibration of the option pricing model.

Analysis of the ill-posedness of the calibrated inverse problems for a time dependent volatility $\sigma(\tau)$ was investigated by many authors. See [35, 52, 55] and references in there.

Uniqueness and stability results for space dependent coefficients $\sigma(y)$ were given by [16]. There, the volatility is considered in Hölder spaces. On of the main results in this issue is

Theorem 1.5.1. *[16, Theorem 1] Let u_1 and u_2 be solutions to the parabolic problem* (9) *and* (10)*, respectively, with $a_1(y)$ and $a_2(y)$ only space dependent, where a_1 and a_2 are Hölder continuous. Moreover, let the corresponding final data given by $u_1^*(y) = u_1(\tau^*, y)$ and $u_2^*(y) = u_2(\tau^*, y)$ and I_0 be a non-empty open subinterval of I.*

If $u_1^*(y) = u_2^*(y)$ on I and $a_1(y) = a_2(y)$ on I_0, then $a_1(y) = a_2(y)$ on I.

If, in addition, $a_1(y) = a_2(y)$ on $I_0 \cup (\mathbb{R} - I)$ and I is bounded, then there exists a constant C that depends only on Hölder norms $|a_1|_\lambda, |a_2|_\lambda, I, I_0, \tau^*$ such that the following *stability estimate holds*

$$|a_1 - a_2|_\lambda \leq C|u_1^* - u_2^*|_{2+\lambda}\,.$$

One of the biggest advantages of considering only space-dependent volatility is that solutions of parabolic problems with a time-independent coefficient are analytic with respect to τ [70].

The ill-posedness of calibrating the local volatility surface $\sigma(T, K)$ is given by

Lemma 1.5.1. *The inverse problem of identifying the local volatility surface $\sigma(T, K)$ from the operator equation* (15) *is ill-posed.*

Proof. The compactness and weak closedness of the operator F, concluded in Theorem 1.3.1, imply the local ill-posedness of the inverse problem of identification of the local volatility surface $\sigma(T, K)$. In fact, for any $\mathcal{H}$-bounded sequence $\{a_n\}_{n\in\mathbb{N}}$ in $\mathcal{D}(F)$, that has no strong convergent subsequences, we can extract an $\mathcal{H}$-weakly convergent subsequence, say $\{a_{n_k}\}_{k\in\mathbb{N}}$. Since $\mathcal{D}(F)$ is weakly closed with respect to the $\mathcal{H}$-norm, the weak limit of $\{a_{n_k}\}_{k\in\mathbb{N}}$ belongs to $\mathcal{D}(F)$. Thus, as F is compact, $\{F(a_{n_k})\}$ has a convergent subsequence, which we again, for simplicity of notation, denote by $\{F(a_{n_k})\}$. So, similar option prices may correspond to completely different volatilities. □

The ill-posedness characteristic of the calibration of the local volatility surface problem calls for regularization methods. One of the contributions of this work is analyzing Tikhonov and iterative regularization to the inverse calibration problem. It is done in Chapters 2 and 3, respectively.

1.6 Review of Volatility Calibration for European Options

In this section, we review some facts about volatility calibration for the standard Black-Scholes model (see [15]) for pricing options. This model assumes that the underlining asset price S satisfies the stochastic differential equation (1). Under a non-arbitrage principle, the European option price U satisfies the parabolic partial differential equation (3) with terminal condition (4).

Specially after the equity market crash in 1987, market participants and academic researchers are all aware that it is impossible to use a single constant σ to fit all European option prices in the market. To overcome those difficulties, some models have been proposed, such as the local volatility function model [33], the stochastic volatility model [56], the jump-diffusion model [9], etc.

The so-called *local volatility* $\sigma(S,T)$ is a quantity of fundamental importance to the trading of options on a stock S. It measures the standard deviation of the rate of change of S.

Obtaining estimates (calibration) for the volatility is a major challenge for market finance. Unlike historical observation of the volatility, based upon a times-series of the stock price, calibration implies an anticipation of the trading agents reflected in the prices of the traded option products derived from S. Because of this, calibration of the volatility σ is crucial to pricing correctly an option.

Volatility calibration (identification) has received intense attention in the last decades. The list of references on the subject is too vast to cite them all here. For some interesting and recent results see, for instance, [7, 16, 22, 23, 33, 35, 52, 54, 55, 59, 68].

One kind of calibration of the volatility σ is estimating it via statistical analysis of historical series for $S = S(t)$. This however dos not lead to the $\sigma(S,T)$ used by market participants to price the options. Another one is to use the so-called *implied volatility.* The implied volatility is obtained by inverting for σ the result of the Black-Scholes formula (1.1). In practice, volatility varies with both strike K and maturity T. This phenomena is known as the volatility *smile.* See [33, 34] for references. It follows that no one constant volatility choice can give prices consistent with the market data.

Assuming that Dupire's model (5) and (6) holds true and that the option prices are known for all possible strikes K and maturities T, the local volatility function $\sigma(T,K)$ can be theoretically found directly from equation (5)

$$\sigma(T,K) = \sqrt{2\left(\frac{U_T + (r-q)KU_K - qU}{K^2 U_{KK}}\right)}. \tag{1.24}$$

In the absence of arbitrage the radicand must be positive. However, this formula has many practical problems. In practice, we have a limited amount of discrete market prices with noise. To make matters worse, the volatility function is extremely sensitive to changes of the prices, i.e., the problem is ill-posed. See Lemma 1.5.1. Hence, the positivity of the radicand or the smoothness restrictions on the volatility function given by (1.24) may be violated. Moreover, there exists numerical instability associated to the ill-posedness of the calibration problem. The reader is invited to review the examples in [23].

Therefore, some regularization strategy is needed for a correct calibration of the local volatility function.

In [68] Lagnardo and Osher proposed a regularized minimization method to fit the smile of option prices using finite difference methods. The minimization process implies a self-stability (regularization) with restriction of σ to the smoothest functional that minimizes the difference between Black-Scholes prices and known market prices. Avellaneda et al. [7] had proposed a representation of the volatility $\sigma(S,t)$ using a relative-entropy minimization method. This method can generate volatilities that change abruptly and drawback numerical solutions. In particular, it may replicate a given set of market prices far from to the numerical solution of Black-Scholes model. In [59] Jackson et al. proposed a spline interpolation so as to choose a smooth volatility function $\sigma(S,t)$. This choice is done via a regularization method that consists in a weighted minimization of the difference between the Black-Scholes prices and the known market prices over a set of strikes and maturities. In [68], a penalization on the smoothness of $\sigma(S,t)$ is added to the minimization functional in [68]. This incorporation implies numerical stability and smoothness of the recovered volatility surface $\sigma(S,t)$. Numerical tests presented on [59] shows that the method is robust. On the other hand, in [59] there is no theoretical proof of efficiency of the proposed method.

Approaches including regularization techniques of the inverse problems have been used to calibrate the local volatility [22, 35, 52, 54, 55]. Instead, there are two standard techniques to solve nonlinear inverse problems. Tikhonov type regularization or iterative type regularization methods [39, 61].

Tikhonov regularization to calibrate volatility model was investigated by [22, 35, 52, 54, 55]. Crépey [22] and after Egger and Engl [35] investigated the stable identification of local volatility surfaces $\sigma(t,S)$ in the Black - Scholes equation from market prices using standard Tikhonov regularization with $\|\cdot\|^2_{\mathcal{H}}$ penalization. Convergence analysis and rates are also discussed in these papers. In [52, 55] the inverse problem of identification of the time dependent volatility function of a European call option with a fixed strike $K > 0$ is considered. In [54], Hofmann *et al.* analyzed the same financial problem of [55] in terms of the Bregman distances applied to a special case where $f(\cdot) = \|\cdot\|^2_{L^2(0,T)}$.

In Chapter 2, we propose Tikhonov regularization by means of a convex regularizing functional as an extension to the quadratic regularization that has been used previously in the inverse problem literature [22, 35, 54]. We address the regularization problem from the perspective of convex analysis methods and Bregman distances. On the theoretical side, our result is that this yields better convergence rates and allows for convergence in spaces different from those in the quadratic regularization setting. In fact, in some cases, the convergence of certain convex regularization expressions implies convergence in the L^1-norm. Moreover, our approach can be applied to some Orlicz-Sobolev norms [1]. Besides those results, our approach connects with central topics in different areas of current research. Such topics include *exponential families* of probability distributions, which is an important subject in Statistics [2, 8].

On the other hand, iterative methods for the regularization of the inverse problems of identification (calibration) the volatility σ is not standard in the literature. As we showed above, this problem typically leads to mathematical models that are ill-posed, i.e., their solution is unstable under data perturbation. Numerical methods that can cope with these problems are the so-called *iterative regularization methods*. In the last years, more emphasis was put on the investigation of iterative regularization methods. It turned out that they are attractive alternatives to Tikhonov regularization, especially for large scale nonlinear inverse problems. The biggest difficulty in applying iterative regularization techniques to nonlinear ill-posed problems comes from the necessary assumptions on the nonlinearity of the problems. It is done by the assumption that the operator F satisfies the so called *local tangential cone condition* (1.21). See [39, 61]. Thanks to Theorem 1.4.2 that verifies the local tangential cone condition (1.21) to the parameter-to-solution map F we can apply iterative regularization to the calibration problem.

Iterative regularization of the local volatility calibration problem is novel to the best of our knowledge. The contribution to the regularization of the local volatility by Landweber iteration is performed in Chapter 3.

Chapter 2

Tikhonov Regularization

The main contribution of the present chapter is that we propose Tikhonov regularization by means of a convex regularizing functional as an extension to the quadratic regularization that has been used previously in the inverse problem literature [7, 22, 35, 54]. We address the regularization problem from the perspective of convex analysis methods and Bregman distances. On the theoretical side, our result is that this yields better convergence rates and allows for convergence in spaces different from those in the quadratic regularization setting as presented in Subsection 2.1.2. Besides those results, our approach connects with central topics in different areas of current research. Such topics include *exponential families* of probability distributions, which is an important subject in Statistics [2, 8]. See Section 2.2.

In general, the theory of regularization of ill-posed problems [10, 13, 39, 79, 81], *Tikhonov Regularization* is a compromise between precision and stability. This compromise is attained from a suitable *a priori* choice of the regularization parameter β in a minimization of the *Tikhonov functional*

$$\mathcal{F}_{\beta,u^\delta}(a) := \frac{1}{2}||F(a) - u^\delta||^2_{L^2(\Omega)} + \beta f_{a_0}(a)\,. \tag{2.1}$$

Minimization of (2.1) is a compromise between minimizing the residual norm (precision) and keeping the "penalization term" f small, i.e., enforcing stability. In general, f takes values on $[0, \infty]$ and a_0 is some *a priori* information about the true solution of the problem. See, for example, [10, 13, 39, 79, 81, 87].

The theory of linear ill-posed problems is by now well understood and pretty complete [10, 39, 64]. Differently from the linear case, where the bound for the operator F is enough to prove Tikhonov regularization, in a nonlinear setting, more restrictive assumptions are needed in order to prove well-posedness of the Tikhonov functional and convergence analysis of the approximated solutions. See, [13, 39, 79].

For fully nonlinear ill-posed problems, the Tikhonov regularization theory is under development. In special, for the last two decades a great deal of articles and books present new results in Tikhonov regularization. See, [13, 39, 79] as some recent references. Using the framework of Bregman distances, new results on convergence analysis and rates to nonlinear inverse problems were derived under generalized source conditions [18, 54, 74, 75, 76].

Motivated by the ideas on [18, 54, 76], in this chapter, we propose Tikhonov regularization by means of a convex regularizing functional as an extension to the quadratic regularization that has been used previously in the inverse problem literature [22, 35]. For instance, it applies to Orlicz-Sobolev regularization [1]. We address the regularization problem from the perspective of convex analysis methods and Bregman distances [18, 54, 74, 76, 79]. On the theoretical side, our result is that this yields better convergence rates and allows for convergence in spaces different from those in the quadratic regularization setting. In fact, in some cases, the convergence of certain convex regularization expressions implies convergence in the L^1-norm.

2.1 Convex Regularization of the Calibration Problem

We apply convex regularization as discussed in [18, 54, 76] to calibrate the local volatility function given by the inverse problem associated to the operator equation (15).

The novelty of the present article *vis-a-vis* [22, 35, 54, 55] is that we consider a regularization method for solving the calibration problem for a general class of convex functionals f. For given a convex f, the proposed method consists in minimizing the Tikhonov functional (2.1) over $\mathcal{D}(F)$, where, $\beta > 0$ is the regularization parameter.

In this paper, we make only the following assumptions on f:

Assumption 2.1.1. *Let $\varepsilon \geq 0$ be fixed. $f : \mathcal{D}(f) \subset \mathcal{H}(\Omega) \longrightarrow [0, \infty]$ is a convex, proper and sequentially weakly lower semi-continuous functional with domain $\mathcal{D}(f)$ containing $\mathcal{D}(F)$.*

An important tool in the studies of Tikhonov type regularization [18, 54, 74, 76] is the

Bregman distance with respect to f.

Definition 2. *Let f as in Assumption 2.1.1. For given $a \in \mathcal{D}(f)$, let $\partial f(a) \subset \mathcal{H}(\Omega)$ denote the subdifferential of the functional f at a, which we define and denote by*

$$\mathcal{D}(\partial f) = \{\tilde{a} : \partial f(\tilde{a}) \neq \emptyset\}$$

the domain of the subdifferential [21]. The Bregman distance with respect to $\zeta \in \partial f(a_1)$ is defined on $\mathcal{D}(f) \times \mathcal{D}(\partial f)$ by

$$D_\zeta(a_2, a_1) = f(a_2) - f(a_1) - \langle \zeta, a_2 - a_1 \rangle \ .$$

Concerning the definition of the subdifferential and the Bregman distance, we emphasize that the subdifferential is a subset of the dual of $\mathcal{H}(\Omega)$. However, in Hilbert spaces there is an isomorphism between the space $\mathcal{H}(\Omega)$ and its dual $\mathcal{H}(\Omega)^*$. This justifies Definition 2 where $\partial f(a)$ is considered a subset of $\mathcal{H}(\Omega)$ and the Bregman distance, which is considered with respect to the $\mathcal{H}(\Omega)$-inner product.

2.1.1 Well-posedness and Convergence Analysis

In the following, we use recent convergence analysis results for nonlinear Tikhonov regularization [79] to prove existence and stability of a regularized solution of (15) by minimization of the Tikhonov functional (2.1). We will first prove that the minimization problem (2.1) has a solution a_β^δ that is stable in the sense of continuous dependence of the solution on the data u^δ. Thus, we make the following abstract assumptions:

Assumption 2.1.2.

1. *The Banach spaces U and V are endowed with topologies τ_U and τ_V that are weaker than the norm topologies. In our context, we later take $U = \mathcal{H}(\Omega)$, $V = L_2(\Omega)$, and endow those spaces with their weak topologies.*

2. *The norm $\|\cdot\|_V$ is sequentially lower semi-continuous with respect to τ_V. In our case V is a Hilbert space and thus the assumption holds.*

3. *The functional $f : \mathcal{D}(f) \subseteq U \to [0, \infty]$ is convex and sequentially lower semi-continuous with respect to τ_U and $\mathcal{D} := \mathcal{D}(F) \cap \mathcal{D}(f) \neq \emptyset$. In the context of this paper, we have $\mathcal{D}(F) \neq \emptyset$ and $\mathcal{D}(F) \subseteq \mathcal{D}(f)$ and thus the assumption is satisfied.*

4. *Let* $\mathcal{F}_{\beta,\bar{u}}$ *the Tikhonov functional defined in* (2.1). *Then,*

$$\mathcal{M}_\beta(M) := level_M(\mathcal{F}_{\beta,\bar{u}}) = \{a : \mathcal{F}_{\beta,\bar{u}}(a) \leq M\}$$

is sequentially pre-compact and closed with respect to τ_U. *The restrictions of* F *to* $\mathcal{M}_\beta(M)$ *are sequentially continuous with respect to the topologies* τ_U *and* τ_V.

The general result of [79] then implies well-posedness, stability, convergence. These results are summarized below.

Theorem 2.1.1 (Existence, Stability, Convergence). *Suppose that* F, f, $\mathcal{D}$, U, *and* V *satisfy Assumption 2.1.2. Furthermore, assume that* $\beta > 0$ *and* $u^\delta \in V$. *Then, we have that*

- *There exists a minimizer of* $\mathcal{F}_{\beta,u^\delta}$.
- *If* (u_k) *is a sequence converging to* u *in* V *with respect to the norm topology, then every sequence* (a_k) *with*

$$a_k \in argmin\big\{\mathcal{F}_{\beta,u_k}(a) : a \in \mathcal{D}\big\}$$

has a subsequence which converges with respect to τ_U. *The limit of every* τ_U*-convergent subsequence* $(a_{k'})$ *of* (a_k) *is a minimizer* $\tilde{a}$ *of* $\mathcal{F}_{\beta,u}$, *and* $\big(f(a_{k'})\big)$ *converges to* $f(\tilde{a})$.

- *If there exists a solution of* (15) *in* $\mathcal{D}$, *then there exists an* f*-minimizing solution of* (15).
- *Assume that* (15) *has a solution in* $\mathcal{D}$ *(which implies the existence of an* f*-minimizing solution) and that* $\beta : (0,\infty) \to (0,\infty)$ *satisfies*

$$\beta(\delta) \to 0 \text{ and } \frac{\delta^2}{\beta(\delta)} \to 0\,, \text{ as } \delta \to 0\,. \tag{2.2}$$

Moreover, we assume that the sequence (δ_k) *converges to* 0, *and that* $u_k := u^{\delta_k}$ *satisfies* $\|\bar{u} - u_k\| \leq \delta_k$.

Set $\beta_k := \beta(\delta_k)$. *Then, every sequence* (a_k) *of elements minimizing* $\mathcal{F}_{\beta_k,u_k}$, *has a subsequence* $(a_{k'})$ *that converges with respect to* τ_U. *The limit* $a^\dagger$ *of any* τ_U *convergent subsequence* $(a_{k'})$ *is an* f*-minimizing solution of* (15), *and* $f(a_k) \to f(a^\dagger)$. *In addition, if the* f*-minimizing solution* $a^\dagger$ *is unique, then* $a_k \to a^\dagger$ *with respect to* τ_U.

The first three conditions on the Assumption 2.1.2 are satisfied for our particular problem. Theorem 1.4.1 implies that Item 4 of Assumption 2.1.2 holds. Therefore, Theorem 2.1.1 is applicable for the functional $\mathcal{F}_{\beta,u^\delta}$ defined in (2.1) to calibrate the local volatility inverse problem.

2.1.2 State of the Art on Convergence Rates

We now summarize the convergence-rate results to the proposed problem available in the literature. In all the examples, the regularization parameter is chosen by $\beta = \beta(\delta) \sim \delta$.

(i) Egger and Engl [35] applied the standard results for nonlinear Tikhonov regularization in a Hilbert space setting, and obtained convergence rates of

$$\left\|a_\beta^\delta - a^\dagger\right\| = \mathcal{O}(\sqrt{\delta}) \qquad \text{and} \qquad \left\|F(a_\beta^\delta) - u^\delta\right\| = \mathcal{O}(\delta) \tag{2.3}$$

to $a_\beta^\delta, a^\dagger \in \mathcal{D}(F) \subset H^1(\Omega)$ under the assumption of the source condition

$$a_0 - a^\dagger = F'(a^\dagger)^* w$$

with $\|w\|$ sufficiently small. Moreover, the above convergence rates are proved for time-independent volatilities in a more regular set and with a variational source condition. See [35, Theorem 4.1].

(ii) Focusing on the time dependent case only, Hofmann and Krämer [55] studied the maximum entropy regularization functional $f(\cdot)$ in the setting of $\mathcal{D}(F) \subset L_1[0,T]$ and data in $L_2[0,T]$. Under the source condition $\log(a^\dagger/\hat{a}) = F'(a^\dagger)^* w$, the rates of

$$\|a_\beta^\delta - a^\dagger\|_{L_1[0,T]} = \mathcal{O}(\sqrt{\delta}) \tag{2.4}$$

were obtained assuming the nonlinear estimate

$$\|F(a) - F(a^\dagger) - F'(a - a^\dagger)\|_{L_2[0,T]} \leq C\|a - a^\dagger\|_{L_1[0,T]}. \tag{2.5}$$

We will return to maximum entropy regularization in Section 2.2 and, more generally, to Bregman distance regularization in Section 2.1.4.

(iii) Hofmann *et al.* [54] improved the convergence rates of [55] for the regularization functional $f(\cdot) = \|\cdot\|_{L_2[0,T]}$. We note that in [54, 55] the volatility parameter is considered to be time-dependent only.

One of the goals of the present subsection is to improve the above convergence rate results by using recent convergence results for the Tikhonov regularization [79]. It is done based on the following theorem which requires further abstract assumptions.

Assumption 2.1.3. *Besides Assumption 2.1.2, we have that*

1. *There exists an f-minimizing solution $a^\dagger$ of (15), which is an element of the Bregman domain $\mathcal{D}_B(f)$.*
2. *There exist $\beta_1 \in [0,1)$, $\beta_2 \geq 0$, and $\xi^\dagger \in \partial f(a^\dagger)$ such that*

$$\langle \xi^\dagger, a^\dagger - a \rangle_{U^*,U} \leq \beta_1 D_{\xi^\dagger}(a, a^\dagger) + \beta_2 \left\| F(a) - F(a^\dagger) \right\| , \tag{2.6}$$

for $a \in \mathcal{M}_{\beta_{max}}(\rho)$, where β_{max}, $\rho > 0$ satisfy the relation $\rho > \beta_{max} f(a^\dagger)$.

Under this assumption we have the following:

Theorem 2.1.2 (Convergence rates [79])**.** *Let F, f, $\mathcal{D}$, U, and V satisfy Assumption 2.1.3. Moreover, let $\beta : (0,\infty) \to (0,\infty)$ satisfy $\beta(\delta) \sim \delta$. Then*

$$D_{\xi^\dagger}(a_\beta^\delta, a^\dagger) = O(\delta) , \qquad \text{and} \qquad \left\| F(a_\beta^\delta) - u^\delta \right\|_V = O(\delta) ,$$

and there exists $c > 0$, such that $f(a_\beta^\delta) \leq f(a^\dagger) + \delta/c$ for every δ with $\beta(\delta) \leq \beta_{max}$.

The following proposition reveals that the technical conditions in Assumption 2.1.2 can be obtained from rather classical ones:

Proposition 2.1.1. *Let F, f, $\mathcal{D}$, U, and V satisfy Assumption 2.1.2. Assume that there exists an f-minimizing solution $a^\dagger$ of (15), and that F is Gâteaux differentiable at $a^\dagger$.*

Moreover, assume that there exists $\gamma \geq 0$ and $\omega^\dagger \in V^*$ with $\gamma \left\| \omega^\dagger \right\| < 1$, such that

$$\xi^\dagger := F'(a^\dagger)^* \omega^\dagger \in \partial f(a^\dagger) \tag{2.7}$$

and there exists $\beta_{max} > 0$ satisfying $\rho > \beta_{max} f(a^\dagger)$ such that

$$\left\| F(a) - F(a^\dagger) - F'(a^\dagger)(a - a^\dagger) \right\| \leq \gamma \, D_{\xi^\dagger}(a, a^\dagger) , \quad a \in \mathcal{M}_{\beta_{max}}(\rho) . \tag{2.8}$$

Then, Assumption 2.1.3 holds.

The convergence rates to the specific inverse problem of calibrate the local volatility require the investigation of (2.6), or alternatively (2.7) and (2.8), respectively. The next subsection is dedicated to the investigation of condition that mimics the above conditions.

2.1.3 Attainment of Source Conditions

The convergence result of Theorem 2.1.2 is directly connected to the existence of a source function w that satisfies the source condition (2.7).

Theorem 2.1.3. *Assume that $\hat{a} \in \mathcal{D}(F) \subset \mathcal{H}(\Omega)$ is a minimizer of (2.1) with u^δ substituted by $\tilde{u}$. Then, there exists $\tilde{w} := \lambda(\tilde{u} - F(\hat{a}))$ such that*

$$\zeta = \lambda F'(\hat{a})^* \tilde{w} \in \partial f(\hat{a})$$

In particular, if $\hat{a} = a^\dagger$, then (2.7) holds.

Proof. The existence of $F'(\hat{a})$ follows from Lemma 1.3.1. Since $\hat{a}$ is a minimizer of (2.1), we must have that [21]

$$0 \in \partial(\|F(\hat{a}) - \tilde{u}\|^2_{L_2(\Omega)} + \beta f(\hat{a})) \subset \partial(\|F(\hat{a}) - \tilde{u}\|^2_{L_2(\Omega)}) + \beta \partial f(\hat{a}))\,. \tag{2.9}$$

Then, if we set $\lambda =: 2/\beta$, it follows from (2.9) that

$$\lambda F'(\hat{a})^*(\tilde{u} - F(\hat{a})) \in \partial f(\hat{a})\,. \tag{2.10}$$

□

We remark that there are examples in linear inverse problem cases where the minimizers of (2.1) are different from $a^\dagger$. See, for example, [39].

It turns out that, for the specific problem under consideration, we are not able to characterize the source condition (2.7). However, we can guarantee (2.6). See section 2.1.4. The first step in order to guarantee (2.6) is the following simple Lemma:

Lemma 2.1.1. *Let $\zeta^\dagger \in \partial f(a^\dagger)$. Then, there exists a function $w^\dagger \in L^2(\Omega)$ and a function $r \in \mathcal{H}(\Omega)$ such that*

$$\zeta^\dagger = F'(a^\dagger)^* w^\dagger + r \tag{2.11}$$

holds. Furthermore, $\|r\|_{\mathcal{H}(\Omega)}$ can be taken arbitrarily small.

Proof. Indeed, Lemma 1.4.1 implies that $\mathcal{R}(F'(a^\dagger)^*)$ is dense in $\mathcal{H}(\Omega)$. □

2.1.4 Convergence Rates

In this subsection, we exhibit a class of functionals for which that we are able to prove that condition (2.6) holds provided the variational source condition (2.11) is satisfied. For that we shall make use of the following concept:

Definition 3. *Let $1 \leq q < \infty$. The Bregman distance $D_\zeta(\cdot, \tilde{a})$ of $f : U \to \mathbb{R} \cup \{+\infty\}$ at $\tilde{a} \in \mathcal{D}_B(f)$ and $\zeta \in \partial f$ is said to be q-coercive with constant $\underline{c} > 0$ if*

$$D_\zeta(a, \tilde{a}) \geq \underline{c} \|a - \tilde{a}\|_U^q, \quad \forall a \in \mathcal{D}(f). \tag{2.12}$$

In the next lemma, we prove that the existence of an approximated source condition as (2.11) and f satisfying Definition 3 is sufficient for convergence rates:

Lemma 2.1.2. *Let $\zeta^\dagger \in \partial f(a^\dagger)$ satisfy (2.11) with $w^\dagger$ and r such that*

$$\underline{c}\big(C\|w^\dagger\|_V + \|r\|_U\big) := \beta_1 \in [0, 1),$$

and the Bregman distance with respect to f is $1-coercive$ with $U := \mathcal{H}(\Omega)$. Then, equation (2.6) holds. In particular, the convergence rates of Theorem 2.1.2 hold.

Proof. Using the continuously Sobolev embedding theorem [1], Equations (2.11) and (1.16), we have that

$$\begin{aligned}
|\langle \zeta^\dagger, a - a^\dagger \rangle| &\leq |\langle \zeta^\dagger - r, a - a^\dagger \rangle + \langle r, a - a^\dagger \rangle| \\
&\leq |\langle w^\dagger, F'(a^\dagger)(a - a^\dagger) \rangle| + \|r\|_U \, \|a - a^\dagger\|_U \\
&\leq \|w^\dagger\|_V \, \|F'(a^\dagger)(a - a^\dagger)\|_V + \|r\|_U \, \|a - a^\dagger\|_U \\
&\leq \big(C\|w^\dagger\|_V + \|r\|_U\big)\|a - a^\dagger\|_U \,.
\end{aligned}$$

From the assumption that f satisfies Definition 3 and the definition of β_1 we have

$$\begin{aligned}
|\langle \zeta^\dagger, a - a^\dagger \rangle| &\leq \big(C\|w^\dagger\|_V + \|r\|_U\big)\|a - a^\dagger\|_U \\
&\leq \beta_1 D_{\zeta^\dagger}(a, a^\dagger) \leq \beta_1 D_{\zeta^\dagger}(a, a^\dagger) + \beta_2 \big\|F(a) - F(a^\dagger)\big\|_V \,.
\end{aligned}$$

The convergence rates now follow from Theorem 2.1.2. □

In the sequel we present possible choices for q-coercive Bregman distance.

Example 2.1.1 (q-coercive Bregman distance)**.** *Let U be a Hilbert space and $\mathcal{D}(f) \subset U$ and $f(a) := q^{-1}\|a - a^\dagger\|_U^q$. Then, the Bregman distance associated to f is q-coercive. See [53] and references therein.*

Example 2.1.2. *Let $1 < q \leq 2$ and $\varepsilon > 0$. We consider the functional*

$$f(a) = \sum_{n=1}^{\infty} | < a, \phi_n > |^q \,,$$

where $\{\phi_n\}$ is an orthonormal basis in $\mathcal{H}(\Omega)$. The functional is convex, proper and sequentially weakly lower semi-continuous. An easy calculation shows that

$$\partial f(a^\dagger) = \sum_{n=1}^{\infty} q| < a^\dagger, \phi_n > |^{q-1}\, sgn(< a^\dagger, \phi_n >)\phi_n \,.$$

Therefore, the Bregman distance of the functional f satisfies

$$\begin{aligned} f(a) - & f(a^\dagger) - \langle \partial f(a^\dagger), a - a^\dagger \rangle \\ & = \sum_{i=1}^{\infty} \left[| < a, \phi_n > |^q - | < a^\dagger, \phi_n > |^q - q| < a^\dagger, \phi_n > |^{q-1}\, sgn(< a^\dagger, \phi_n >)\langle \phi_n, a - a^\dagger \rangle \right] \\ & \geq C_q \sum_{n=1}^{\infty} |\langle a - a^\dagger, \phi_n \rangle|^2 = C_q \|a - a^\dagger\|_U^2 \,. \end{aligned}$$

Here, we use the following inequality [32, Section 5]

$$C_q|x - y|^2 \leq (|y|^2 - |x|^2 - q|x|^{q-1}\, sgn(x)(y - x))$$

in the estimate. Hence, f is 2-coercive. According to Lemma 2.1.2 and Equation (2.13) the rate of $\mathcal{O}(\delta)$ holds for the $\mathcal{H}(\Omega)$-norm.

This method is usually considered in the case of sparsity regularization [47]. The case $p = 1$, which refers to the original sparsity regularization is not taken into account here, since we aim at convergence rates in the Hilbert's space norm.

Other interesting possibilities would be to consider looking for Orlicz-Sobolev norms [1].

With the assumption of Lemma 2.1.2, if in addition f is q-coercive, a convergence rate in the norm holds:

$$\left\|a_\beta^\delta - a^\dagger\right\|_U^q \leq D_{\zeta^\dagger}(a_\beta^\delta, a^\dagger) = \mathcal{O}(\delta) \ , \quad 1 \leq q < \infty \, . \tag{2.13}$$

In other words, convergence of q-coercive Bregman distances imply in convergence in the norm.

2.2 Exponential Families

The concept of exponential family arises naturally in order to answer the question: What is the maximum entropy distribution consistent with given constraints on expected values?

Given the interpretation of option prices as expected values, with respect to appropriate measures, which depend on the local volatility surface, a minute's thought leads us to naturally associate the problem of volatility estimation from observed option prices to exponential measures. Financially, it can be understood as follows: Each volatility surface leads to a corresponding risk neutral measure whose expectation of the payoff is the observed derivative prices. Thus, if we are given the problem of inferring the volatility surface from the market observed option prices, the use of Bregman distances leads to the choice of certain exponential families of probability distributions. The latter, can be thought of as optimal (in an appropriate sense) *a posteriori* distributions for the class of models under consideration. This hints to yet another connection with the now classical work developed by Avellaneda *et al.* See [7] and references therein.

Because of its minimum information, entropy regularization has been proposed by several authors [37, 40, 67, 75]. The Kullback-Leibler-fuctional (see (2.16) below) allows one to perform regularization from an information perspective, in the sense that one constraint, the closeness of the data a to its observed perturbation a^δ to satisfy an information measure rather than some distance measure associated with some functional space. In this way, a more natural characterization of the relationship between the data and its observational realization is achieved, and thereby becomes a basis for assessing the convergence os the regularized solution to the exact data. Moreover, The Kullback-Leibler-fuctional has numerous properties that can be exploited mathematically [19, 75, 74].

In this section, we will motivate the use of Bregman distances for regularization from a statistical perspective and then connect it to the general theory developed earlier.

The Darmois-Koopman-Pitman theorem states that under certain regularity conditions on the probability density, a necessary and sufficient condition for the existence of a sufficient statistic of fixed dimension is that the probability density belongs to the exponential family [4]. Formally, a sufficient statistic is a statistic for which the conditional distribution of the original data, given that statistic, depends only on that data. This leads naturally to the idea that the value of any definition of information must, with respect to the set of sufficient statistic transformations that can be applied to the given measurements, be invariant over the corresponding set of transformed measurements [67].

We start with the definition of an *exponential family* in dimension 1, which is used later on to define appropriate priors.

Definition 2.2.1 (Regular Exponential Family). *Let $\psi : \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$ be convex and $p_0 : \mathbb{R} \to \mathbb{R}_+$ by continuous. The family of functions $p_{\psi,\theta} : \mathbb{R} \to \mathbb{R}_+$ defined by*

$$p_{\psi,\theta}(s) := \exp(s \cdot \theta - \psi(\theta))p_0(s)$$

is called a regular exponential family. In this context the function ψ is called log-partition *or* circulant *function. The* expectation number $a(\theta)$ *is defined by*

$$a(\theta) := \int_{\mathbb{R}} s p_{\psi,\theta}(s)\, ds\ .$$

This definition calls for an example, namely:

Example 2.2.1. *We consider the exponential family of normal distributions on $\mathbb{R}$ with known variance $\varpi^2 = 1$. The density is*

$$p_{\psi,\theta}(s) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(s-\theta)^2}{2}\right)\ , \quad s > 0.$$

This is a one parameter exponential family with

$$p_0(s) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{s^2}{2}\right) \ \textit{and}\ \psi(\theta) = \frac{\theta^2}{2}\ ,$$

The expectation number is

$$
\begin{aligned}
a(\theta) &= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} s \exp\left(-\frac{(s-\theta)^2}{2}\right) ds \\
&= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} (s-\theta) \exp\left(-\frac{(s-\theta)^2}{2}\right) ds + \frac{\theta}{\sqrt{2\pi}} \int_{\mathbb{R}} \exp\left(-\frac{s^2}{2}\right) ds \\
&= 0 + \theta \, .
\end{aligned}
$$

We have the following result from [8] which relates exponential families with Bregman distances.

Theorem 2.2.1. *[8, Theorem 4] Let ψ^* denote the Fenchel transform of ψ, which we assume to be differentiable. Then, the Bregman distance with respect to ψ^* is given by*

$$D_{\psi^*}(\hat{a}, \tilde{a}) = \psi^*(\hat{a}) - \psi^*(\tilde{a}) - \psi^{*\prime}(\tilde{a})(\hat{a} - \tilde{a}) \, .$$

We assume that $a(\theta) \in int(dom(\psi^*))$. Then,

$$p_{\psi,\theta}(a) = \exp\big(- D_{\psi^*}(a, a(\theta)) \big) \exp\big(\psi^*(a)\big) p_0(a) \, . \tag{2.14}$$

An interesting example is given by

Example 2.2.2 (Exponential Families and Fenchel conjugate)**.** *For a Gaussian distribution $\psi(\theta) = \frac{\varpi^2}{2}\theta^2$, then $\psi^*(a) = \frac{a^2}{2\varpi^2}$. For Poisson distribution $\psi(\theta) = \exp(\theta)$ we have $\psi^*(a) = a\log(a) - a$.*

2.2.1 Bregman Distance Regularization

Among others, Bregman distance becomes, for appropriate choices of a generating functional, the Hilbert space norm and the Kulback-Leibler distance. The generality of this framework for the analysis of entropy regularization and its extensions have proven successful in establishing error estimates and convergence rates [18, 54, 76, 74].

We shall now motivate Bregman distance regularization as a log-maximum a-posteriori estimator for an exponential family. The connection between Bregman distances and exponential families is well established in some contexts [2, 8], albeit in the present context our motivation in this section is heuristic. From the financial intuition, it can be understood as follows: Each volatility surface leads to a corresponding risk neutral measure whose expectation of the payoff is the observed derivative prices. Thus, if we are given the problem of inferring the volatility surface from market observed option prices, the use of Bregman distances leads to the choice of certain exponential families of probability distributions. The latter, can be thought of as optimal (in an appropriate sense) *a posteriori* distribution for the class of models under consideration. Indeed, under some circumstances, exponential families are connected to minimal entropy measures. This hints to yet another connection with the now classical work developed by Avellaneda *et al.* See [7] and references therein.

For the moment, for motivation purposes, we consider a discrete statistical setting. As usual, we denote by $(\mathcal{X}, \mathcal{F}, \mathbb{P})$ a probability space. We let $\vec{x} := (x_i)_i$ be a sequence of elements in $\mathcal{X}$ and $\vec{a} = (a_i)_i$, where $a_i = a(x_i) \in \mathbb{R}$. We assume that the conditional probability density for observable data $u_i^\delta := u^\delta(x_i)$ from $u_i := F(a)(x_i)$ are normally and identically distributed with mean zero and variance ϖ^2. That is, the probability of observing u_i^δ given u_i is given by

$$p(u_i^\delta | u_i) = \frac{1}{\varpi\sqrt{2\pi}} \exp\left(-\frac{|u_i^\delta - u_i|^2}{2\varpi^2} \right) .$$

Now, for $a \in \mathbb{R}_{>0}$ denote $\theta := \theta(a)$. With this notation, for some prior $\hat{a}$, the *a priori* distribution is defined by

$$p(a) := p_{\psi,\theta}(\hat{a}) = \exp\left(\hat{a}\theta - \psi(\theta)\right) p_0(\hat{a}) \ .$$

In order to clarify this formula, recall that θ depends on a and this is the only a dependence that shows up on the right hand side.

This in turn, according to Theorem 2.2.1, can be rewritten as

$$p(a) = \exp\left(-D_{\psi^*}(\hat{a}, a)\right) \exp(\psi^*(\hat{a})) p_0(\hat{a}) \ \ .$$

The advantage of this representation is that it does not involve any parametrization of the exponential family (that is, with respect to θ). In this context the Log-maximum estimation then consists in minimizing the functional

$$\vec{a} \longmapsto \sum_i \big(-\log(p(u_i^\delta|u_i)) - \log(p(a_i)) \big) \ ,$$

which is equivalent to minimizing the functional

$$\vec{a} \longmapsto \sum_i (u_i - u_i^\delta)^2 + \beta \sum_i D_{\psi^*}(\hat{a}_i, a_i) \ ,$$

where $\beta = 2\varpi^2$. Note that the Bregman distance is in general not symmetric, and we minimize with respect to the second component of the Bregman distance.

In summary, we have shown that Bregman distance regularization can be considered a log maximum a-posteriori estimator for the expectation number, in our case for the expected volatility.

In this model, we shall introduce some regularization techniques. For notational simplicity, we formulate them in an infinite dimensional framework. Hereafter, we shall assume again that Ω is a bounded subdomain of $\mathbb{R}^2$. With this framework, we remark that $\mathcal{D}(F) \subset \mathcal{H}(\Omega) \cap L^\infty_{>0}(\Omega) \subset L^1(\Omega)$, where $L^\infty_{>0}(\Omega)$ is the set of functions that are (essentially) bounded from below and above by some positive constants.

Example 2.2.3. *According to Example 2.2.2, if we use the exponential family associated to Poisson distributions, we obtain Kullback-Leibler regularization, consisting in minimization of*

$$a \longmapsto \mathcal{F}_{\beta,u^\delta}(a) := \left\| F(a) - u^\delta \right\|^2_{L^2(\Omega)} + \beta KL(\hat{a}, a) \ , \tag{2.15}$$

where

$$KL(\hat{a}, a) = \int_\Omega a \log\left((\hat{a}/a) - (\hat{a} - a)\right) \, dx \ . \tag{2.16}$$

We note that the Kullback-Leibler distance is the Bregman distance associated to the Boltzmann-Shannon entropy

$$\mathcal{G}(a) := \int_\Omega a \log(a) \, dx \ . \tag{2.17}$$

We also note that the standard Kullback-Leibler regularization [75], and more generally, the Bregman distance regularization, is in general considered with respect to the first component.

However, the modelling with exponential families results in Bregman distances with respect to the second component.

Remark 2.2.1. *The domains of $\mathcal{G}$, $\mathcal{D}(\mathcal{G})$, and of the subgradient of $\mathcal{G}$, $\mathcal{D}(\partial\mathcal{G})$, are $L^\infty_{\geq 0}(\Omega)$ (the set of bounded non-negative functions) and $L^\infty_{>0}(\Omega)$, respectively.*

The Kullback-Leibler distance, which is the Bregman distance of the Boltzmann-Shannon entropy, is defined the Bregamn domain on $\mathcal{D}_B(\mathcal{G})$, that is a subset of $L^\infty_{>0}$. Moreover, the Kullback-Leibler distance is lower semi-continuous with respect to the L^1-norm [75]. Based on this property we extend the Kullback-Leibler distance, to take value $+\infty$ if either $a \notin \mathcal{D}(\mathcal{G})$ or $b \notin \mathcal{D}_B(\mathcal{G})$.

Note that there are exceptional cases, when the integral

$$\int_\Omega (a\log(a/\hat{a}) - (a - \hat{a}))\ dx$$

is actually finite, but $KL(a, \hat{a}) = \infty$. *This can be seen by taking for instance* $a \in L^1_{>0}(\Omega)$ *which is not in* $L^\infty(\Omega)$ *and* $\hat{a} = Ca$, *where* C *is a constant. The reason here, is that* a *is not an element of the subgradient of the Boltzmann-Shannon entropy.*

This follows directly from the definition of the domains of the convex functionals and subgradients.

Example 2.2.4. *Alternatively to the Example 2.2.3, if we use the exponential family associated to Gaussian distributions in the Example 2.2.2, we obtain the standard quadratic Tikhonov functional, consisting in*

$$a \longmapsto \mathcal{F}_{\beta,u^\delta}(a) := \left\|F(a) - u^\delta\right\|^2_{L^2(\Omega)} + \beta\|a - \hat{a}\|^2 \ . \tag{2.18}$$

2.2.2 Convergence Analysis

To prove that the minimization of $\mathcal{F}_{\beta,u^\delta}$ defined in (2.15) is well–posed, we have to choose appropriate spaces and topologies first. We choose $\tau_{\tilde{U}}$, $\tau_{\tilde{V}}$ the weak topologies on $L^1(\Omega)$ and $L^2(\Omega)$, respectively.

As it follows, we have some auxiliary lemmas from [37, 75].

Lemma 2.2.1. *Let* $a, \hat{a} \in \mathcal{D}(\partial\mathcal{G})$ *as in the Remark 2.2.1. Then, the following statements holds:*

(i) The function $(a, \hat{a}) \longmapsto KL(a, \hat{a})$ *is convex.*

(ii) For any $C > 0$ *and any* $\hat{a} \in L^\infty_+(\Omega)$ *the sets*

$$\mathcal{E}_C := \{a \in L^1(\Omega)\ :\ KL(a, \hat{a}) \le C\,\}$$

are weakly closed in $L^1(\Omega)$.

(iii) The functional $a \longmapsto KL(a, \hat{a})$ *is weakly lower semicontinuous in the* $\tau_{\tilde{U}}$*-topology.*

(iv) The sets $\mathcal{E}_C$ *are weakly compact subset of* $L^1(\Omega)$.

Lemma 2.2.2. *Let Ω be a bounded subset of $\mathbb{R}^2$ with Lipschitz boundary. Moreover, assume that F is continuous with respect to the weak topologies on $L^1(\Omega)$ and $L^2(\Omega)$, respectively.*

1. *Let $a, b \in \mathcal{D}(\mathcal{G})$. Then*

$$\|a - b\|^2_{L^1(\Omega)} \leq \left(\frac{2}{3}\|a\|_{L^1(\Omega)} + \frac{4}{3}\|b\|_{L^1(\Omega)}\right) KL(a, b) \; . \tag{2.19}$$

   Here, we use the convention that $0 \cdot (+\infty) = 0$.

2. *With the generalization of the Kullback-Leibler distance. For sequences $(a_k)_k$ and $(b_k)_k$ in $L^1(\Omega)$, such that one of them is bounded: If the sequence $KL(a_k, b_k) \to 0$, then $\|a_k - b_k\|_{L^1(\Omega)} \to 0$.*

3. *Let $0 \neq \hat{a} \in \mathcal{D}_B(\mathcal{G})$, then the sets*

$$\mathcal{M}_{\beta,u^\delta}(M) := \{a \in \mathcal{D}_B(\mathcal{G}) : \mathcal{F}_{\beta,u^\delta}(a) \leq M\}$$

   are $\tau_{\tilde{U}}$ sequentially compact.

Proof. For the proofs of Item 1 and Item 2 see [75]. To prove Item 3, we use (2.19). Let $(a^k)_k$ be a sequence in $\mathcal{M}_{\beta,u^\delta}(M)$, then according to (2.19), it is uniformly bounded in $L^1(\Omega)$. According to Lemma 2.2.1 the KL functional satisfies

$$\mathrm{KL}(\hat{a}, \tilde{a}) \leq \liminf \mathrm{KL}(\hat{a}, a^k)$$

Now, since F is continuous with respect to weak topologies on $L^1(\Omega)$ and $L^2(\Omega)$, it follows that

$$\left\|F(\tilde{a}) - u^\delta\right\|^2_{L^2(\Omega)} + \beta \mathrm{KL}(\hat{a}, \tilde{a}) \leq M \; .$$

□

Note that (2.19) implies that convergence of a sequence $\{a_n\}$ to some a in the KL-distance, together with convergence in the weak topology implies in the strong convergence in $L^1(\Omega)$.

Definition 4. *We say that a minimizer of $\mathcal{F}_{\beta,u^\delta}$ as defined in the Equation (2.15) is stable if the argmin $\mathcal{F}_{\beta,u^{\delta_k}} \to$ argmin $\mathcal{F}_{\beta,u^0}$ for $\delta_k \to 0$ and that argmin $\mathcal{F}_{\beta(\delta_k),u^{\delta_k}}$ converges to a solution of (15) with minimal energy.*

The above Lemmas 2.2.1 and 2.2.2 imply immediately in the well posedness and convergence to the functional (2.15). Indeed, we have

Theorem 2.2.2. *Let Ω bounded. There exists a minimizer of $\mathcal{F}_{\beta,u^\delta}$ in (2.15). The minimizers are stable and convergent for $\beta(\delta) \to 0$ and $\delta^2/\beta(\delta) \to 0$.*

Proof. Note that, since $\mathcal{D}(F) \subset \mathcal{H}(\Omega)$ we have $\mathcal{D}(F) \subset \mathcal{D}_B(\mathcal{G})$. From Theorem 1.3.1, $F : \mathcal{D}(F) \subset \mathcal{H}(\Omega) \to W_2^{1,2}(\Omega)$ is weakly continuous. Since Ω is bounded, we have $\mathcal{D}(F) \subset \mathcal{H}(\Omega) \subset L^2(\Omega) \subset L^1(\Omega)$ and $W_2^{1,2}(\Omega) \subset L^2(\Omega)$, with continuous embedding. It follows that $F : \mathcal{D}(F) \subset L^1(\Omega) \to L^2(\Omega)$ is weakly continuous, i.e., it satisfies the assumptions on the Lemma 2.2.2. Moreover, $\mathcal{D}(F)$ is a convex and closed subset of $\mathcal{E}_C$ for some $C > 0$ sufficiently large. Hence, it is weakly closed on the $L^1(\Omega)$-topology. The prove follows the standard arguments on Tikhonov regularization of nonlinear inverse problems [39, 79] together with the results on Lemmas 2.2.1 and 2.2.2. □

2.2.3 Convergence Rates

An important consequence of (2.19) and Theorem 2.1.2 is that, with the framework developed in Section 2.2, is possible to get convergence rates in the $L^1(\Omega)$. In particular, this extends the convergence rates results obtained by [55] to the local volatility calibration framework.

Theorem 2.2.3. *Given the assumptions in Section 2.2, we have the following rates of convergence for the regularized solution obtained by minimization of the Tikhonov functional* (2.15)

$$\left\| a_\beta^\delta - a^\dagger \right\|_{L^1(\Omega)} = \mathcal{O}(\sqrt{\delta})\,. \tag{2.20}$$

Proof. Follows directly from (2.19) and the convergence rates in Theorem 2.1.2 to the Bregman distance. □

Therefore, let δ_k be a sequence converging to zero and $a_k = a_{\beta_k}^{\delta_k}$ the respective minimizers of the Tikhonov functional (2.1). Take the sequence $b_k = a^\dagger$ for all $k \in \mathbb{N}$. Then, from Lemma 2.2.2 Item 2 we have

$$\left\| a_k - a^\dagger \right\|_{L^1(\Omega)} \to 0\,, \quad as \quad \delta_k \to 0\,.$$

Chapter 3

Iterative Regularization

The main contribution of this chapter is the verification and implementation of iterative regularization to the local volatility calibration problem. The main result of Subsection 3.1.1 is the Lemma 3.1.1. In the Section 3.2 we extend the framework of stability and convergence of the Landweber iteration to the $L^2(\Omega)$-norm by means of a non-classical discrepancy principle. A numerical implementation that confirms the robustness of the Landweber iteration for the calibration problem is presented in Section 3.3. Moreover, derivation of Kaczmarz type regularization methods are presented in Section 3.4.

The minimization of the Tikhonov functional is often performed by means of iterative methods in order to solve the first order necessary conditions on the minimizer. Historically, iterative methods had been used to solve well-posed problems. However, many iterative methods also have interesting self regularization properties if one terminates the iteration early enough [39, 61]. Thus, it allows iterative methods to be successfully applied to solve many inverse problems. An example of this idea is employed in many algorithms of Computerized Tomography [72]. Differently from Tikhonov regularization [13, 39, 79, 87], in classical iterative methods [13, 39, 51, 61, 69] the index of the iteration plays the role of the regularization parameter, and the stopping index of the iteration plays the role of parameter selection method [39, 61].

Many iterative methods are based on solving the first-order optimality condition for the least square problem

$$\min \quad \frac{c}{2}\|F(a) - u\|^2 , \quad a \in \mathcal{D}(F) , \tag{3.1}$$

i.e., solving the normal equations

$$0 = cF'(a)^*(u - F(a))\,. \tag{3.2}$$

Thus, to solve (3.2), the iteration is defined as an appropriate fixed point equation given by

$$a = \psi(a) := a + cF'(a)^*(u - F(a))\,, \tag{3.3}$$

where $F : \mathbb{D}(F) \subset \mathbb{X} \longrightarrow \mathbb{Y}$ is assumed to be Fréchet differentiable and $\mathbb{X}$ and $\mathbb{Y}$ are Hilbert spaces.

Convergence of (3.3), for well-posed problems, is typically obtained by contractive (or nonexpansive) operators ψ. The constant c is a weight used to guarantee the nonexpansivity of ψ. On the other hand, in the case of an ill-posed problem, to check analytically whether the operator ψ is contractive is almost impossible in many practical examples. Therefore, in iterative methods for nonlinear inverse problems, the contractiveness of ψ is replaced by local properties, as the *tangential cone condition* and the local scaling of the Fréchet derivative of F (see Theorem 1.4.2 and Equation (1.15)). Both properties are strong enough to ensure the local convergence of some iterative regularization methods to a solution of (15). Moreover, it guarantees that the iteration remains in $\mathcal{D}(F)$, which makes the iteration well defined.

Our interest here is to analyze iterative regularization theory applied to the inverse problem (15), with application to the calibration of the local volatility surface in the Black-Scholes model (9) and (10). In this chapter, we focus on the Landweber iteration [39, 61, 69]. Other methods, like the steepest descent, iterated Tikhonov and Newton type methods to solve ill-posed problems have similar ideas associated. Convergence analysis of these iterative methods is well known [14, 39, 51, 61, 78]. Of course, each of these methods requires different assumptions in order to guarantee convergence of the iteration.

3.1 The Classical Landweber Regularization

In 1951, Landweber [69] proved strong convergence of the method of successive approximations applied to (3.3) for a linear compact operator F. The generalization of the Landweber idea to nonlinear operators [39, 61] consists in iterating Equation (3.3) as

$$a_{k+1}^{\delta} = a_k^{\delta} + cF'(a_k^{\delta})^*(u^{\delta} - F(a_k^{\delta}))\,. \tag{3.4}$$

In a noisy case, the iterative methods (3.4) only becomes a regularization method, if an early stop of the iterative process is considered. In other words, only if, for a suitable stopping index k_*, the iterate $a_{k_*}^{\delta}$ yields a stable approximation to the solution a^* of the operator equation (15). Differently from the Tikhonov regularization, the stopping index of the iteration plays the role of the regularization parameter. A successful early stopping strategy is determined by the stopping index $k_*(\delta, y^{\delta})$ given by the *discrepancy principle* (see [39, 61])

$$\left\| u^{\delta} - F(a_{k_*(\delta,y^{\delta})}^{\delta}) \right\| \leq r\delta < \left\| u^{\delta} - F(a_k^{\delta}) \right\|, \tag{3.5}$$

where

$$r > 2\frac{1+\eta}{1-2\eta}\,, \tag{3.6}$$

is a relaxation term. By this, we mean that, if the iteration is stopped at index $k_*(\delta, y^{\delta})$ such that for the first time, the residual becomes small compared to the quantity $r\delta$.

The convergence analysis for the Landweber iteration (3.4), for a nonlinear inverse problem, in general, does not hold globally [61]. As in the linear case [39], the Landweber iteration (3.4) only converges if

$$\|F'(a)\| \leq 1/c \qquad a \in B_{2\rho}(a_0) \subset \mathbb{D}(F)\,, \tag{3.7}$$

where $F'(\cdot)$ is the Fréchet derivative of the operator F defined in (15). This is the motivation of adding a relaxation parameter c at iteration (3.4).

To prove local convergence we shall use the following assumption:

Assumption 3.1.1. *We will assume that the equation $F(a) = u(a) - u(a_0)$, as in* (15)*, is solvable in the open ball $B_\rho(a_0) \subset \mathcal{D}(F)$, where the operator $F : \mathcal{D}(F) \subset \mathbb{X} \longrightarrow \mathbb{Y}$ is continuous and weakly closed with a continuous Fréchet derivative $F'(\cdot)$.*

The assumption of solvability of (15), in the open ball $B_\rho(a_0) \subset \mathcal{D}(F)$, implies the existence of a solution of minimal distance to a_0. The so called a_0-*minimum-norm* solution. It is, often, denoted by $a^\dagger$. In general, uniqueness of $a^\dagger$ does not hold. However, we have the following:

Proposition 3.1.1. *[61, Proposition 2.1] Let $\rho, \varepsilon > 0$ be such that*

$$\|F(a) - F(\tilde{a}) - F'(a)(a - \tilde{a})\| \leq c(a, \tilde{a})\|F(a) - F(\tilde{a})\| \,, \quad a, \tilde{a} \in B_\rho(a_0) \subset \mathcal{D}(F) \,, \tag{3.8}$$

for some $c(a, \tilde{a}) > 0$, where $c(a, \tilde{a}) < 1$ if $\|a - \tilde{a}\| \leq \varepsilon$.
If $F(a) = u(a) - u(a_0)$ is solvable in $B_\rho(a_0)$, then a unique a_0-minimal-norm solution exists and satisfies the condition

$$a^\dagger - a_0 \in \mathcal{N}(F'(a^\dagger))^\perp \,. \tag{3.9}$$

The local convergence of the iteration (3.4) is summarized in the following result:

Theorem 3.1.1. *Assume that Equations (3.7) and (1.21) are satisfied. Consider the Landweber iteration (3.4). For any $a^* \in B_\rho(a_0)$, a solution of (15), and a noise level $\left\|u^\delta - u^*\right\| \leq \delta$, we have that*

1. *While $\left\|u^\delta - F(a_k^\delta)\right\| \geq r\delta$*

$$\left\|a^* - a_{k+1}^\delta\right\| \leq \left\|a^* - a_k^\delta\right\| \,. \tag{3.10}$$

Moreover, if $a_0 \in B_\rho(a^*) \subset \mathcal{D}(F)$, then $a_k^\delta \in B_{2\rho}(a^*)$ for all k and

$$k_*(r\delta)^2 \leq \sum_{k=0}^{k_*-1} \left\|u^\delta - F(a_k^\delta)\right\|^2 \leq \frac{r\|a^* - a_0\|^2}{(1-2\eta)r - 2(1+\eta)} \,, \quad \forall\, 0 \leq k \leq k_* \,. \tag{3.11}$$

In particular, if $u^\delta = u$ (i.e., $\delta = 0$), then

$$\sum_{k=0}^{\infty} \|u - F(a_k)\|^2 < \infty \,. \tag{3.12}$$

2. *If there exists $a^* \in B_\rho(a_0)$ solution of (15) and $\delta = 0$, then the iterated a_k given by (3.4) converges to a^*.*

3. *In the noisy data case, if the iterations are stopped according to the discrepancy principle (3.5) and r is given by (3.6), then $a_{k(\delta, u^\delta)}^\delta$ converges to a solution of (15) as $\delta \to 0$.*

Proof. See, for example, the Chapter 2 of [61] . □

Details of the proof of Theorem 3.1.1 follow the similar analysis done in the Section 3.2 below.

In contrast to Tikhonov regularization, the existence of a *source condition* [39, 61] is not enough to prove convergence rates for the Landweber iteration. Indeed, in iterative methods, the Fréchet derivative $F'(\cdot)$ changes for each iteration, with the iterate a_k. This implies that assumptions on source conditions are more restrictive than in Tikhonov-type methods [39, 61]. Moreover, even if we can prove convergence rates for the Landweber iteration, the convergence is, in general, arbitrarily slow. For results on convergence rates of iterative methods and respective source condition see [39, 61].

3.1.1 Application to the Local Volatility Calibration Inverse Problem

The application of the Landweber iteration to calibrate the local volatility requires that the operator F, defined in Equation (15), satisfies the assumption of Theorem 3.1.1.

Lemma 3.1.1. *Let $\mathcal{D}(F) \subset \mathcal{H}(\Omega)$. The operator $F : \mathcal{D}(F) \subset \mathcal{H}(\Omega) \longrightarrow W_2^{1,2}(\Omega)$, defined in Equation* (15)*, satisfies the assumptions of Theorem 3.1.1. Moreover, there exists a unique a_0-minimal-norm solution.*

Proof. It follows from Theorem 1.3.1 that F is compact, weakly (sequentially) continuous and weakly closed. Lemma 1.3.1 implies that the operator F has a continuous and uniformly bounded Fréchet derivative, with $\|F'(\cdot)\| \leq C$. Moreover, from Theorem 1.4.2, the operator F satisfies the local tangential cone condition (1.21). The existence of a unique a_0-minimal-norm-solution follows from (1.23) in Theorem 1.4.2 and Proposition 3.1.1. □

Lemma 3.1.1 tells us that, we can apply the Landweber iteration (3.4) to the volatility calibration inverse problem.

We stress that Theorem 3.1.1 is valid with the discrepancy principle measured in $W_2^{1,2}(\Omega)$-norm. In this framework, the implementation of the Landweber iteration (3.4) requires us to calculate $F'(\cdot)^*$ in the $W_2^{1,2}$-inner product. However, this is not an easy task.

In order to simplify the calculation of $F'(\cdot)^*$, we look for the convergence of the Landweber iteration (3.4) in the L^2-norm. With this new framework, $F'(\cdot)^*$ can be calculated in the $L^2(\Omega)$-inner product. See Section 3.2 below.

3.2 Landweber Iteration with a Non-Classical Discrepancy Principle

The convergence analysis of the Landweber iteration in Section 3.1 is based on the discrepancy principle (3.5) measured with the $W_2^{1,2}(\Omega)$-norm. This means that, for each iteration, we have to compare the residual norm in $W_2^{1,2}(\Omega)$. This becomes numerically expensive and difficult to implement, since, the adjoint of $F'(\cdot)$ in the $W_2^{1,2}$-inner product is not so easy to be calculated.

In this section, we plan to recover the theoretical assertions of Theorem 3.1.1, with the discrepancy principle available in the $L^2(\Omega)$-norm. The principal improvement in the analysis of the iteration (3.4) in the L^2-norm is an obvious simplification in the numerical implementation of the algorithm. See Subsection 3.3.

We start with the convergence analysis to the Landweber iteration in the L^2-norm by stating some auxiliary lemmas. They characterize the necessary assumptions for the convergence of the iteration in the L^2-norm.

Lemma 3.2.1. *The operator $F : \mathcal{D}(F) \subset \mathcal{H}(\Omega) \longrightarrow L^2(\Omega)$ is continuous, compact, weakly (sequentially) closed and Fréchet differentiable. The Fréchet derivative satisfies*

$$\|F'(a)h\|_{L^2(\Omega)} \leq C_1 \|h\|_{\mathcal{H}(\Omega)}\,, \qquad \forall a \in B_{2\rho}(a_0) \subset \mathcal{D}(F)\,. \tag{3.13}$$

Moreover, the local tangential cone condition (1.21) *can be replaced by*

$$\|F(a) - F(\tilde{a}) - F'(\tilde{a})(a - \tilde{a})\|_{L^2(\Omega)} \leq \eta \|F(a) - F(\tilde{a})\|_{W_2^{1,2}(\Omega)}\,, \tag{3.14}$$

for all $a, \tilde{a} \in B_{2\rho}(a_0) \subset \mathcal{D}(F)$ and with $\eta < \frac{1}{2}$.

Proof. This follows from the continuous embedding of $W_2^{1,2}(\Omega)$ in $L^2(\Omega)$ applied to (1.16) in Lemma 1.3.1 and (1.21) in Theorem 1.4.2. □

3.2.1 Convergence Analysis

Before starting with the convergence analysis of the nonlinear Landweber iteration (3.4), we would like to emphasize that for a fixed index k the iterated a_k^δ depends continuously on the data u^δ. This is the case because a_k^δ is a combination of a chain of continuous operations.

In order to provide regularization properties of the nonlinear Landweber iteration with the discrepancy principle in $L^2(\Omega)$, we assume that $u^\delta \in W_2^{1,2}(\Omega)$ and satisfies

$$\left\|u^* - u^\delta\right\|_{W_2^{1,2}(\Omega)} \leq \delta\,. \tag{3.15}$$

The intuition behind the data measurements in $W_2^{1,2}(\Omega)$ is that, we can have access to the information about the solution of (9) and its derivatives in $L^2(\Omega)$. We recognize that this is a restrictive assumption, since, in practice, the data are measured in $L^2(\Omega)$. On the other hand, given the data as in (3.15), we can measure the discrepancy principle in the $L^2(\Omega)$-norm.

We start by an auxiliary lemma.

Lemma 3.2.2. *Assume that the conditions* (3.13) *and* (3.14) *hold for all* $a, \tilde{a} \in B_{2\rho}(a_0)$ *with* $C_1 \leq 1$. *Moreover, assume that the equation* $F(a) = u$ *has a solution* $a^\dagger \in B_\rho(a_0)$. *If* $a_k^\delta \in B_\rho(a^\dagger)$, *then, while*

$$\left\|u^\delta - F(a_k^\delta)\right\|_{L^2(\Omega)} > 2\frac{1+\eta}{1-2\eta}\delta \tag{3.16}$$

we have

$$\left\|a_{k+1}^\delta - a^\dagger\right\|_{L^2(\Omega)}^2 - \left\|a_k^\delta - a^\dagger\right\|_{L^2(\Omega)}^2 \leq \left\|u^\delta - F(a_k^\delta)\right\|_{L^2(\Omega)} \left[2\frac{1+\eta}{1-2\eta}\delta - \left\|u^\delta - F(a_k^\delta)\right\|_{L^2(\Omega)}\right] \tag{3.17}$$

Proof. Let $a_k^\delta \in B_\rho(a^\dagger)$. It is follows from the triangular inequality that $a^\dagger, a_k^\delta \in B_{2\rho}(a_0)$.

Hence, (3.13) and (3.14) are applicable. From (3.4) and (3.15) we obtain

$$
\begin{aligned}
&\|a_{k+1}^\delta - a^\dagger\|_{L^2(\Omega)}^2 - \|a_k^\delta - a^\dagger\|_{L^2(\Omega)}^2 = 2\langle a_{k+1}^\delta - a_k^\delta, a_k^\delta - a^\dagger\rangle + \|a_{k+1}^\delta - a_k^\delta\|_{L^2(\Omega)}^2 \\
&= 2\langle F'(a_k^\delta)^*(u^\delta - F(a_k^\delta)), a_k^\delta - a^*\rangle + \|F'(a_k^\delta)^*(u^\delta - F(a_k^\delta))\|_{L^2(\Omega)}^2 \\
&\le 2\langle u^\delta - F(a_k^\delta), -F'(a_k^\delta)(a^* - a_k^\delta)\rangle + \|u^\delta - F(a_k^\delta)\|_{L^2(\Omega)}^2 \\
&= 2\langle u^\delta - F(a_k^\delta), u^\delta - F(a_k^\delta) - F'(a_k^\delta)(a^* - a_k^\delta)\rangle - 2\|u^\delta - F(a_k^\delta)\|_{L^2(\Omega)}^2 + \|u^\delta - F(a_k^\delta)\|_{L^2(\Omega)}^2 \\
&\le 2\|u^\delta - F(a_k^\delta)\|_{L^2(\Omega)}\|u^\delta - F(a_k^\delta) - F'(a_k^\delta)(a^* - a_k^\delta)\|_{L^2(\Omega)} - \|u^\delta - F(a_k^\delta)\|_{L^2(\Omega)}^2 \\
&\le \|u^\delta - F(a_k^\delta)\|_{L^2(\Omega)}\left[2\delta + 2\|F(a^*) - F(a_k^\delta) - F'(a_k^\delta)(a^* - a_k^\delta)\|_{L^2(\Omega)} - \|u^\delta - F(a_k^\delta)\|_{L^2(\Omega)}\right] \\
&\le \|u^\delta - F(a_k^\delta)\|_{L^2(\Omega)}\left[2\delta + 2\eta\|F(a^*) - F(a_k^\delta)\|_{W_2^{1,2}(\Omega)} - \|u^\delta - F(a_k^\delta)\|_{L^2(\Omega)}\right] \\
&\le \|u^\delta - F(a_k^\delta)\|_{L^2(\Omega)}\left[2(1+\eta)\delta + 2\eta\|u^\delta - F(a_k^\delta)\|_{W_2^{1,2}(\Omega)} - \|u^\delta - F(a_k^\delta)\|_{L^2(\Omega)}\right]
\end{aligned}
$$

From the continuous embedding of $W_2^{1,2}(\Omega)$ in $L^2(\Omega)$ and (3.16) we have

$$
\begin{aligned}
&\|a_{k+1}^\delta - a^\dagger\|_{L^2(\Omega)}^2 - \|a_k^\delta - a^\dagger\|_{L^2(\Omega)}^2 \\
&\qquad \le \|u^\delta - F(a_k^\delta)\|_{L^2(\Omega)}\left[2(1+\eta)\delta + 4\eta\frac{1+\eta}{1-2\eta}\delta - \|u^\delta - F(a_k^\delta)\|_{L^2(\Omega)}\right] \\
&\qquad = \|u^\delta - F(a_k^\delta)\|_{L^2(\Omega)}\left[2\frac{1+\eta}{1-2\eta}\delta - \|u^\delta - F(a_k^\delta)\|_{L^2(\Omega)}\right]
\end{aligned}
$$

□

The following Proposition has a key meaning in the convergence analysis of Landweber iterations.

Proposition 3.2.1 (Monotonicity). *Assume that assumptions of Lemma 3.2.2 hold. Then, a_{k+1}^δ is a better approximation of $a^\dagger$ than a_k^δ. Moreover, $a_k^\delta, a_{k+1}^\delta \in B_\rho(a^\dagger)$.*

Proof. The first assertion follows, since, the right hand side of (3.17) is negative. The second assertion follows by an inductive argument as Theorem 3.1.1. □

Given the proof of Lemma 3.2.2, we are able to guarantee that the stopping index k_* in (3.5), available in the $L^2(\Omega)$-norm, is finite and hence, well defined.

Corollary 3.2.1. *Assume that the assumptions of Lemma 3.2.2 hold. Let k_* be chosen according to the stopping rule* (3.5) *in the $L^2(\Omega)$-norm, with r given by* (3.6). *If $\delta > 0$ then, k_* is finite.*

In particular, if $\delta = 0$ then

$$\sum_{k=0}^{\infty} \|u - F(a_k)\|^2_{L^2(\Omega)} < \infty \,. \tag{3.18}$$

Proof. Given that $a_0^\delta = a_0 \in B_\rho(a^\dagger)$, it follows from induction and (3.17) that $a_k^\delta \in B_\rho(a^\dagger)$ for all $0 \le k < k_*$. Hence, Proposition 3.2.1 is applicable for all $0 \le k < k_*$. Adding up (3.17) with k from 0 through $k_* - 1$, we obtain

$$\sum_{k=0}^{K_*-1} \left\|u^\delta - F(a_k^\delta)\right\|_{L^2(\Omega)} \left[\left\|u^\delta - F(a_k^\delta)\right\|_{L^2(\Omega)} - 2\frac{1+\eta}{1-2\eta}\delta\right]$$
$$\le \|a_0 - a^*\|^2_{L^2(\Omega)} - \left\|a^\delta_{k_*} - a^\dagger\right\|^2_{L^2(\Omega)} \le \left\|a_0 - a^\dagger\right\|^2_{L^2(\Omega)} .$$

Using (3.5) and (3.6), we have that

$$k_* r\delta^2 \left(r - 2\frac{1+\eta}{1-2\eta}\right) \le \sum_{k=0}^{K_*-1} \left\|u^\delta - F(a_k^\delta)\right\|_{L^2(\Omega)} \left[\left\|u^\delta - F(a_k^\delta)\right\|_{L^2(\Omega)} - 2\frac{1+\eta}{1-2\eta}\delta\right] . \tag{3.19}$$

Given the choice of r as in (3.6), it follows that the left hand side of (3.19) is positive. Hence, k_* is finite. Obviously, if $\delta = 0$, then Proposition 3.2.1 follows for all k. From (3.19), we have (3.18). □

Note that (3.18) is also followed by the continuous embedding of $W_2^{1,2}(\Omega)$ in $L^2(\Omega)$ and Theorem 3.1.1. Equation (3.18) means that, if the Landweber iteration is running with noise-free data ($\delta = 0$) and if the iteration converges, then the limit is a solution of $F(a) = u$. The convergence of the Landweber iteration with accurate data will be proved in the next theorem. The proof of this theorem follows similar arguments to those found in [61]. However, it differs in a few arguments. Therefore we prefer to present the proof in a full version.

Theorem 3.2.1 (Convergence for exact data)**.** *Suppose the assumptions of Lemma 3.2.2 are satisfied. Then, the Landweber iteration* (3.4)*, applied to exact data u, converges to an a_0-minimum-norm solution $a^\dagger$ of $F(a) = u$.*

Proof. The existence of a unique a_0-minimum-norm solution $a^\dagger$ follows from Lemma 3.1.1. Let

$$e_k = a_k - a^\dagger \,.$$

Then, Proposition 3.2.1 implies that $\{\|e_k\|\}$ is monotonically decreasing. Hence, it converges to some $\varepsilon > 0$. The convergence argument is based in showing that $\{e_k\}$ is a Cauchy sequence. Given $j \geq k$, select an index j, with $k \leq l \leq j$, such that

$$\left\|F(a^\dagger) - F(a_l)\right\|_{L^2(\Omega)} \leq \left\|F(a^\dagger) - F(a_i)\right\|_{L^2(\Omega)} \qquad \text{for all} \qquad k \leq i \leq j\,. \tag{3.20}$$

The triangular inequality implies that

$$\|a_j - a_k\|_{L^2(\Omega)} \leq \|a_j - a_l\|_{L^2(\Omega)} + \|a_l - a_k\|_{L^2(\Omega)}\,.$$

We are going to show that $\|a_j - a_l\|_{L^2(\Omega)}$ converges to zero as $k \to \infty$. A similar argument applied to $\|a_l - a_k\|_{L^2(\Omega)}$ concludes the assertion. Note that

$$\|a_j - a_l\|^2_{L^2(\Omega)} = 2\langle a_j - a_l, e_l\rangle + \|a_j\|^2_{L^2(\Omega)} - \|a_l\|^2_{L^2(\Omega)}\,,$$

and the last two terms on the right hand side go to $\varepsilon^2 - \varepsilon^2 = 0$ for $k \to \infty$. Moreover, applying (3.14)

$$\begin{aligned}
|\langle a_j - a_l, e_l\rangle| &= \left|\sum_{i=l}^{j-1} \langle F'(a_i)^*(u - F(a_i)), e_l\rangle\right| \leq \sum_{i=l}^{j-1} \left|\langle u - F(a_i), F'(a_i)(a_l - a^\dagger)\rangle\right| \\
&\leq \sum_{i=l}^{j-1} \left\|F(a^\dagger) - F(a_i)\right\|_{L^2(\Omega)} \left\|F'(a_i)(a_l - a_i + a_i - a^\dagger)\right\|_{L^2(\Omega)} \\
&\leq \sum_{i=l}^{j-1} \left\|F(a^\dagger) - F(a_i)\right\|_{L^2(\Omega)} \big(\left\|F(a_i) - F(a^\dagger) - F'(a_i)(a^\dagger - a_i)\right\|_{L^2(\Omega)} \\
&\qquad + \left\|F(a^\dagger) - F(a_l)\right\|_{L^2(\Omega)} + \|F(a_i) - F(a_l) - F'(a_i)(a_l - a_i)\|\big) \\
&\leq \sum_{i=l}^{j-1} \left\|F(a^\dagger) - F(a_i)\right\|_{L^2(\Omega)} \Big(\eta \left\|F(a^\dagger) - F(a_i)\right\|_{W_2^{1,2}(\Omega)} \\
&\qquad + \left\|F(a^\dagger) - F(a_l)\right\|_{L^2(\Omega)} + \eta \|F(a_l) - F(a_i)\|_{W_2^{1,2}(\Omega)}\Big) \\
&\leq (1 + 3\eta) \sum_{i=l}^{j-1} \left\|F(a^\dagger) - F(a_i)\right\|^2_{W_2^{1,2}(\Omega)}\,,
\end{aligned}$$

where to obtain the last inequality, we used the continuous embedding of $W_2^{1,2}(\Omega)$ in $L^2(\Omega)$ and equation (3.20). From Theorem 3.1.1, the sequence $\{\|u - F(a_i)\|^2_{W_2^{1,2}(\Omega)}\}$ converges to zero as $k \to \infty$. Hence, $\{e_k\}$ and consequently $\{a_k\}$ are Cauchy sequences. Now, equation

(3.18) implies that the limit of a_k should be a solution of $F(a) = u$.

By the definition of the Landweber iteration (3.4) and Lemma 1.4.1

$$a_{k+1} - a_k \in \mathcal{R}(F'(a_k)^*) \subset \mathcal{N}(F'(a_k))^\perp \subset \mathcal{N}(F'(a^\dagger))^\perp$$

and hence

$$a_k - a_0 \in \mathcal{N}(F'(a^\dagger))^\perp \qquad \text{for all} \qquad k \in \mathbb{N}\,. \tag{3.21}$$

Hence the limit of $\{a_k\}$ also satisfies (3.21). As $a^\dagger$ is the unique solution for which (3.21) holds, it follows that $a_k \to a^\dagger$. □

The next theorem shows that the stopping rules given by (3.5), in the $L^2(\Omega)$-norm, renders the Landweber iteration a regularization method.

Theorem 3.2.2. *Suppose the assumptions of Theorem 3.2.1 hold and $k_* = k_*(\delta, u^\delta)$ is selected according to the stopping rule* (3.5) *in the $L^2(\Omega)$-norm. Then, the Landweber iteration $a^\delta_{k_*}$ converges to $a^\dagger$ as $\delta \to 0$.*

Proof. See [61]. □

3.3 Numerical Validation

Successful numerical strategy for pricing options, for example, binomial trees, Monte Carlo methods, finite-element and finite difference methods are frequently used. See references and implementations in [23, 34, 35, 59].

The numerical implementation of the nonlinear Landweber iteration requires the evaluation of the gradient direction $F'(a_k)^*(u^\delta - F(a_k))$. In the calibration problem, this means that, for each step of the algorithm, we need to solve the PDE (9) and evaluate the adjoint operator $F'(a_k)$. Depending of the inner product, the numerical evaluation of $F'(\cdot)^*$ would be expensive.

However, in the $L^2(\Omega)$-inner product, the evaluation of the iteration (3.4) can be performed efficiently via a variational approach.

Denote by $R = u^\delta - u(a)$. Let W the solution of the adjoint equation

$$W_\tau + (aW)_{yy} + (aW)_y = R\,, \tag{3.22}$$

with homogeneous boundary and final condition. Then, formally, we have

$$\langle R, u'(a)[h]\rangle_{L^2(\Omega)} = \int_\Omega (u'(a)[h])(\tau, y)R(\tau, y)dyd\tau \tag{3.23}$$
$$= \int_\Omega (u'(a)[h])(W_\tau + (aW)_{yy} + (aW)_y)dyds$$
$$= \int_\Omega [-(u'(a)[h])_\tau + a((u'(a)[h])_{yy} - (u'(a)[h])_y)W dyds$$
$$= \int_\Omega -h(u_{yy} - u_y)W dyds\,. \tag{3.24}$$

Summarizing, each step of the Landweber iteration requires a solution of parabolic PDE (9) and (3.22), with respective boundary and initial conditions. Moreover, the evaluation of (3.23).

Below we present some examples of reconstructions using the nonlinear Landweber iteration (3.4).

The routines to solve (9) and the adjoint equation (3.22) are implemented using Crank-Nicholson methods [83] with an explicit drift term. In other words, we use a forward-difference approximation to the time partial derivative to obtain the explicit scheme

$$\frac{u_n^{m+1} - u_n^m}{\delta\tau} + \mathcal{O}(\delta\tau) = \frac{u_{n+1}^m - 2u_n^m + u_{n-1}^m}{(\delta y)^2} + \mathcal{O}((\delta y)^2)\,,$$

and a backward difference to obtain an implicit scheme

$$\frac{u_n^{m+1} - u_n^m}{\delta\tau} + \mathcal{O}(\delta\tau) = \frac{u_{n+1}^{m+1} - 2u_n^{m+1} + u_{n-1}^{m+1}}{(\delta y)^2} + \mathcal{O}((\delta y)^2)\,.$$

The average of these two equations is

$$\frac{u_n^{m+1} - u_n^m}{\delta\tau} + \mathcal{O}(\delta\tau) = \frac{1}{2}\left(\frac{u_{n+1}^m - 2u_n^m + u_{n-1}^m}{(\delta y)^2} + \frac{u_{n+1}^{m+1} - 2u_n^{m+1} + u_{n-1}^{m+1}}{(\delta y)^2}\right) + \mathcal{O}((\delta y)^2)\,.$$

Ignoring the error terms leads to the Crank-Nicholson methods

$$u_n^{m+1} - \frac{\alpha}{2}\left(u_{n+1}^{m+1} - 2u_n^{m+1} + u_{n-1}^{m+1}\right) = u_n^m + \frac{\alpha}{2}\left(u_{n+1}^m - 2u_n^m + u_{n-1}^m\right)\,,$$

where $\alpha := \frac{\delta\tau}{(\delta y)^2}$. Here, $\delta\tau$ and δy represent the time and space step size, respectively. Hence $\alpha < 1$ for the stability condition to be satisfied.

Again, we assume that we can truncate the infinite mesh $y = N^+\delta y$ and $y = N^-\delta y$, and take N^+ and N^- sufficient large, such that no errors are introduced. This restriction implies in artificial boundary conditions $u^m_{N^-}$ and $u^m_{N^+}$.

The algorithm stars with $u_n^{m+1} = u_n^m$. Then, we shift to the right using the boundary condition $u^m_{N^-}$ to evaluate the left hand side of the Crank-Nicholson methods and shift to the left using the boundary condition $u^m_{N^+}$ to evaluate the right hand side of the Crank-Nicholson methods. The internal looping is implemented using SSOR iteration without relaxation.

Examples

For the numerical implementation, the domain needs to be restricted to a finite region. In our examples, we solve the equation (9) with initial condition (10) and the adjoint equation (3.22) on the domain $\Omega_0 := (0, 0.2) \times (-5, 5)$. This restriction to the domain Ω_0 implies the introduction of artificial boundary condition $u(-5, \tau) = 1$ and $u(5, \tau) = 0$. With this choice, it is possible to prove that the error introduced is approximately 10^{-3}.

In the present validation we use as data generator the same code that produces the reconstruction. Thus we are committing so-called "inverse crime" [39]. Moreover, in all examples, we take the initial guess a_0 identically 1.

Example 3.3.1. *[Noise free data] In the first example, Figure 3.1 the value of the exact variance a that solves the inverse problem. Figures 3.2 and 3.3 show the reconstructed parameters obtained from the Landweber iteration. Figure 3.5 shows the respective error between the true parameter and the calculated, without noise in the data. The numer-*

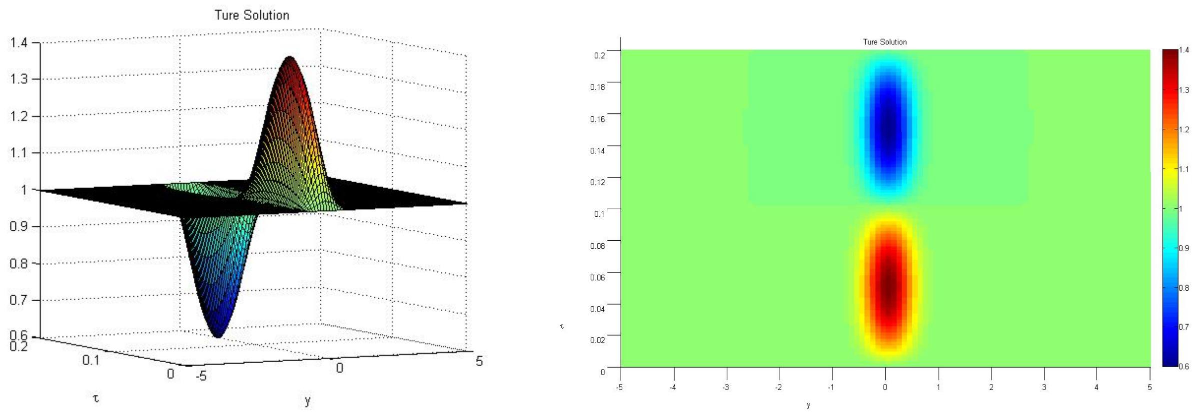

Figure 3.1: Exact variance. Left, described as a surface. Right, as an intensity plot.

ical results in Example 3.3.1 shows that, in the noise free case, the Landweber iteration

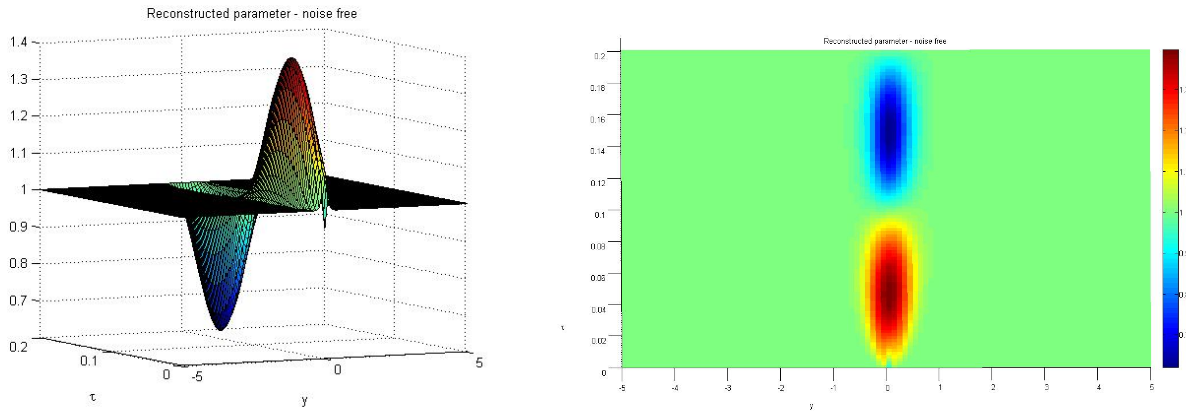

Figure 3.2: Reconstructed parameter after 500 iteration steps. Noise free data.

converges to a solution of the problem (15)*, as* $k \to \infty$*. See Figures 3.4. This confirms numerically the results in Theorem 3.1.1.*

Example 3.3.2. *[*0.5% *of noise in the data] In this example, we have* 0.5% *of noise in the data in the* $L^2(\Omega)$*-norm (or* 0.1% *in the* $L^\infty(\Omega)$*-norm as shows in Figure 3.6), concentrated in a region of the form* $y \in [-5/3, 5/3]$ *and* $\tau \in [1/30, 1/5]$ *as shown in Figure 3.6. As before, Figure 3.1 shows the exact solution of the inverse problem. Figure 3.7 shows the reconstructed solution using the Landweber iteration. Figure 3.8, shows the error between the true parameter and the calculated one as the iterations progress. The Landweber iteration stops after 1800 steps, as shown in Figure 3.9, confirms the discrepancy principle, with* $r = 4.8$*. Example 3.3.2 shows that, in the noisy case, the calculated solution is not as good as in the noise free case. On the other hand, Figures 3.7 and 3.8 validate, after the discrepancy principle, the result in Theorem 3.1.1.*

Example 3.3.3. *[*5% *of noise in the data] In this example, we have* 5% *of noise in the data, measured in the* $L^2(\Omega_0)$*-norm (or* 1% *of noise in the* $L^\infty(\Omega_0)$*-norm) concentrated in a region of the form* $y \in [-5/3, 5/3]$ *and* $\tau \in [1/30, 1/5]$ *as shown in Figure 3.10. As before, Figure 3.1 shows the exact solution. Figure 3.11 shows the reconstructed solution using the Landweber iteration. Figure 3.12, shows the error between the true variance and the computed one. The Landweber iteration stops after 500 steps as shown in Figure 3.13, confirming the discrepancy principle, with* $r = 4.8$*. In the bottom right corner of Figure 3.12 we show the reconstructed parameter after 2000 iterations. In Figure 3.13 we also plot the evolution of the residual and the error before the discrepancy principle, i.e., after 2000 iteration. The result shows that the residual of the iteration is monotone, but the* *iteration is not [61]. Hence, we should have stopped the iteration in accordance with the discrepancy principle.*

The financial interpretation of the choice of domains in Example 3.3.2 and 3.3.3 is the following: The market data is given for a finite set of maturities and strikes. Moreover, prices are influenced by the bid-ask spreads. Therefore, the actual set of observed prices, for the calibration problem, are only known up to some noise level δ.

Example 3.3.4. *[Noise free data] In this example, we have the true parameter given by Figure 3.14. Note that, the structure of the solution is more complicated. Figure 3.15 shows the approximated solution after* 4000 *steps of the Landweber iteration.*

In Example 3.3.4, we see that, after 4000 iterations, the error between the true solution and the reconstructed is bigger than the same approximation given in Example 3.3.1. This phenomenon is expected, since, the initial guess in both examples is the same. Hence, the initial guess is far from the true solution in Example 3.3.4, compared with the situation in Example 3.3.1. However, we have good numerical approximation result in Example 3.3.4.

A final remark about the numerical implementation is that the inverse problem is hard to solve. Given our choice of $\tau = 0.2$, $y \in (-5, 5)$ and the mesh with $\delta\tau = 0.002$ and $\delta y = 0.1$, the matrix representing u has dimension 101×101. The examples are obtained in an Intel Core2 with 1.5 Hz. The total processing time for Example 3.3.1 was approximately 2 minutes.

3.4 Kaczmarz Strategies Applied to the Inverse Problem of Option Prices

In this section, we analyze Kaczmarz type strategies [11, 24, 50, 49, 66] applied to the calibration problem for European call option prices.

The motivation is the following: In practice, options are only sold for a very few maturities, typically, only one pre-assigned day per month qualifies for a maturity time. We denote them by $T_1, \cdots, T_N$. For each T_j, $j = 1, \cdots N$, a certain number of strikes is available, and we shall assume, for simplicity, that the smallest and largest of these, K_{min} and K_{max} are the same for each maturity. Given the above motivation, the inverse problem of option price seeks for σ given

$$U^j(t, S, K, T_j) = U^j_*(K)\,, \qquad k \in [K_{min}, K_{max}]\,. \tag{3.25}$$

Here S is the market price of the stock at time t, and $U_*^j(K)$ denotes the market price of options with different strikes for a correspondent maturity T_j, for $j = 1, \cdots, N$.

As before, we know that the option premium $U_j(\cdot, \cdot, K, T_j)$ satisfies the Dupire's equation (9) with the initial value (10), for each T_j.

For simplicity, we assume that $\sigma = \sigma(K)$ and that $b = q - r$ is constant, i.e., is independent of the time. In this framework, the nonlinear inverse problem is the stable calibration of the volatility $\sigma(K)$ by observation of market prices (3.25), associated with the solutions of the Dupire's equation (5) with initial condition (6), to match quoted market prices $U_*^j(K)$, for a given set of maturities T_j, with $j = 1, \cdots, N$.

Making use of the change of variables (7) and (8), we obtain a system of PDEs

$$- u_\tau^j + a(y)(u_{yy}^j - u_y^j) + b\, u_y^j = 0 \text{ in } \Omega_j \tag{3.26}$$

$$u^j(0, y) = S_0(1 - e^y)^+ , \tag{3.27}$$

where, $\Omega_j = (0, T_j) \times I$, for $j = 1, \cdots, N$. Here, I denotes the interval $[K_{min}, K_{max}]$ in the y-variable.

In the analysis that follows, we are interested in a continuous observation of prices $U_*^j(K)$ as in (3.25). Since, financial markets typically allow only a few and prefixed maturity dates with a discrete sample of strikes for each maturity and, the prices are defined as bid-ask spreads, this leads to an interpolation procedure or approximation to the value of the option. Therefore, the actual set of observed prices U^{j,δ_j}, or input data, for the calibration, is only known up to some noise level δ_j, which we assume to satisfy

$$\left\| \bar{u}^j - u^{j,\delta_j} \right\| \leq \delta_j \, . \tag{3.28}$$

Thus, the inverse problem can now be formulated as follows: Calibrate the volatility parameter $a(y)$ in the system of nonlinear operator equations

$$F_j(a) = u^j \, , \qquad j = 1, \cdots, N \, , \tag{3.29}$$

where $F_j : \mathcal{D}(F_j) \subset \mathcal{H}(\Omega) \longrightarrow W_2^{1,2}(\Omega_j)$ is the parameter-to-solution map and $u^j = u^j(a)$ is the solution of (9) and (10) to the corresponding T_j.

Here, we assume that the admissible set of parameters $\mathcal{D}(F_j)$ is defined by

$$\mathcal{D}(F_j) := \{a \in \mathcal{D}(F) \, : \, a = a(y) \text{ and } a_0 \text{ is constant}\}, \quad \forall j \in \{1, \cdots, N\} \, .$$

Remark 3.4.1. *Note that, all the assumptions of Theorem 3.1.1 are satisfied by operators* F_j *on* (3.29)*, for any* $j \in \{1, \cdots, N\}$*.*

3.4.1 Kaczmarz Type Strategies Applied to Solving Nonlinear System of Equations.

Standard methods for the solution of the system (3.29) are based on the use of Tikhonov type regularization methods [13, 39, 79] or iterative type regularization methods [39, 61] after rewriting (3.29) as a single equation $F(a) = u(a)$, where

$$F := (F_1, \cdots, F_N) : \bigcap_{j=1}^{N} \mathcal{D}(F_j) \longrightarrow (W_2^{1,2}(\Omega_j))^N \,,$$

with $u(a) = (u^1(a), \cdots, u^N(a))$. However, these methods become inefficient and expensive if N is large. In such situation, Kaczmarz type methods [60] are more efficient and are often used in practice. A group of N subsequent steps (starting at some multiple k of N) shall be called a cycle. Kaczmarz type algorithm, consist in, given a starting point a_0, cyclically consider each equation in (3.29), separately. A famous Kaczmarz type algorithm is the ART iteration in Computerized Tomography [72].

Recently, Kaczmarz type methods for systems of ill-posed equations were analyzed. We refer the reader to [11, 12, 17, 24, 49, 50, 66].

In [66] the *Landweber-Kaczmarz method*(LK) that consists in applying the Landweber iteration (3.4) to each equation on (3.29) as a cycle was analyzed. The authors used the termination of the iteration as the first index such that the discrepancy principle (3.5) is attained for *any* component of the system (3.29). In [49, 50] the Landweber-Kaczmarz approach was analyzed with the incorporation of a bang-bang relaxation parameter in the classical Landweber-Kaczmarz iteration [66], combined with a new stopping rule. This is called *loping-Landweber-Kaczmarz iteration*(lLK). Indeed, it consists on the iteration

$$a_{k+1}^{\delta} = a_k^{\delta} + \omega_k F'_{[k]}(a_k^{\delta})(u_{[k]}^{\delta} - F_{[k]}(a_k^{\delta})) \tag{3.30}$$

with

$$\omega_{[k]} = \omega_{[k]}(\delta, u_{[k]}^{\delta}) = \begin{cases} 1 & \text{if } \left\| u_{[k]}^{\delta} - F_{[k]}(a_k^{\delta}) \right\| > \tau\delta \\ 0 & \text{otherwise}\,. \end{cases} \tag{3.31}$$

Here, $[k] = k \, mod \, N \in \{0, \cdots, N-1\}$. The iteration is stopped when $\omega_{[k]} = 0$ in a cycle.

In [24] the loped strategy was analyzed together with the steepest-descent iteration to solve nonlinear systems of ill-posed equations. It is called the *loping-Steepest-Descent-Kaczmarz* (l-SDK). The (I-SDK) iteration consists in incorporating the loped parameter $\omega_{[k]}$ in the steepest-descent iteration coupled with a Kaczmarz strategy. Once again, the iteration is stopped when $\omega_{[k]} = 0$ in a cycle. In [12] the so-called *loping-Levenberg-Marquardt-Kaczmarz*(l-LMK) iteration is proposed to solve (3.29). This corresponds to introducing the loped parameter $\omega_{[k]}$ in the Levenberg-Marquardt iteration and use a cycle strategy. Finally, in [11] *iterated-Tikhonov-Kaczmarz*(iTK) and *loping-iterate-Tikhonov-Kaczmarz*(L-iTK) methods are analyzed. This consists in incorporating the Kaczmarz and loped strategy in the iterated Tikhonov method. The difference with respect to the above methods is that these are implicit methods (see [11]).

The regularization properties of the Kaczmarz type strategies are summarized in the next Theorem.

Theorem 3.4.1. *Let $F_j : \mathcal{D}(F_j) \subset \mathbb{X} \longrightarrow \mathbb{Y}_j$ satisfying all the assumption on Theorem 3.1.1, for all $j \in \{1, \cdots, N\}$. Let the stopping index k_* be such that $\omega_{[k_*]} = 0$ in a cycle. Then, Kaczmarz type strategies (LK), (l-LK), (l-SDK) and (l-LMK) are regularization methods for solving nonlinear systems of equations* (3.29). *Moreover, if the Fréchet derivative of F_j is locally Lipschitz continuous, then the iteration (iTK) and (L-iTK) are also regularization methods.*

Proof. See in [66, 50, 24, 12, 11]. □

Corollary 3.4.1. *The Kaczmarz strategies in the Theorem 3.4.1 can be applied to the calibration problem in* (3.29).

Proof. Indeed, the system of nonlinear operators (3.29) satisfies all the assumption of Theorem 3.4.1 as we show below.

First, by definition, $\mathcal{D}(F_j) \subset \mathcal{D}(F) \subset \mathcal{H}(\Omega)$. Hence, from Theorem 1.3.1 and Lemma 1.3.1, we have that F_j is continuous and weakly closed and with a continuous Fréchet derivative, for $j = 1, \cdot, N$. Moreover, Lemma 1.3.1 implies that the Fréchet derivative of F_j is locally Lipschitz continuous, for $j = 1, \cdot, N$.

Now, Theorem (1.4.2) applied to each F_j, for $j = 1, \cdot, N$, guarantees that the local tangential cone condition (1.21) is satisfied. □

Numerical implementation and comparison of Kaczmarz type strategies to the calibration problem will be performed in future works.

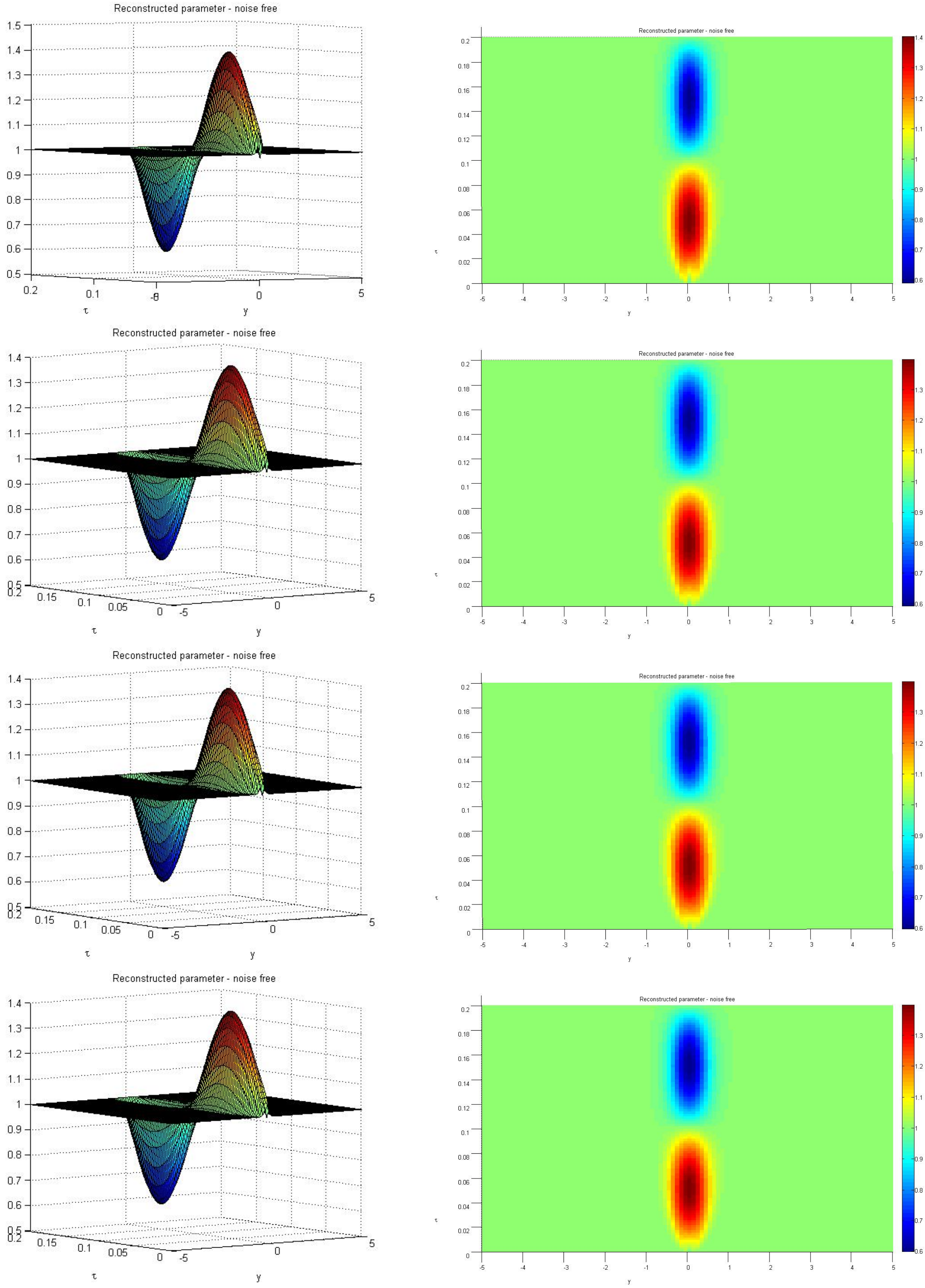

Figure 3.3: Reconstructed variance after 1000, 2000, 3000 and 4000 iteration steps respectively. Noise free data.

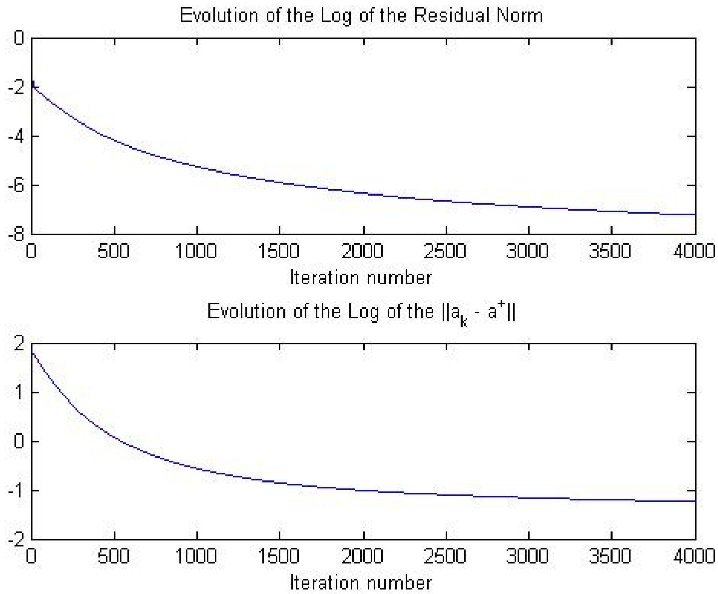

Figure 3.4: Evolution of the log of the residual and error in the approximate $L^2(\Omega)$-norm.

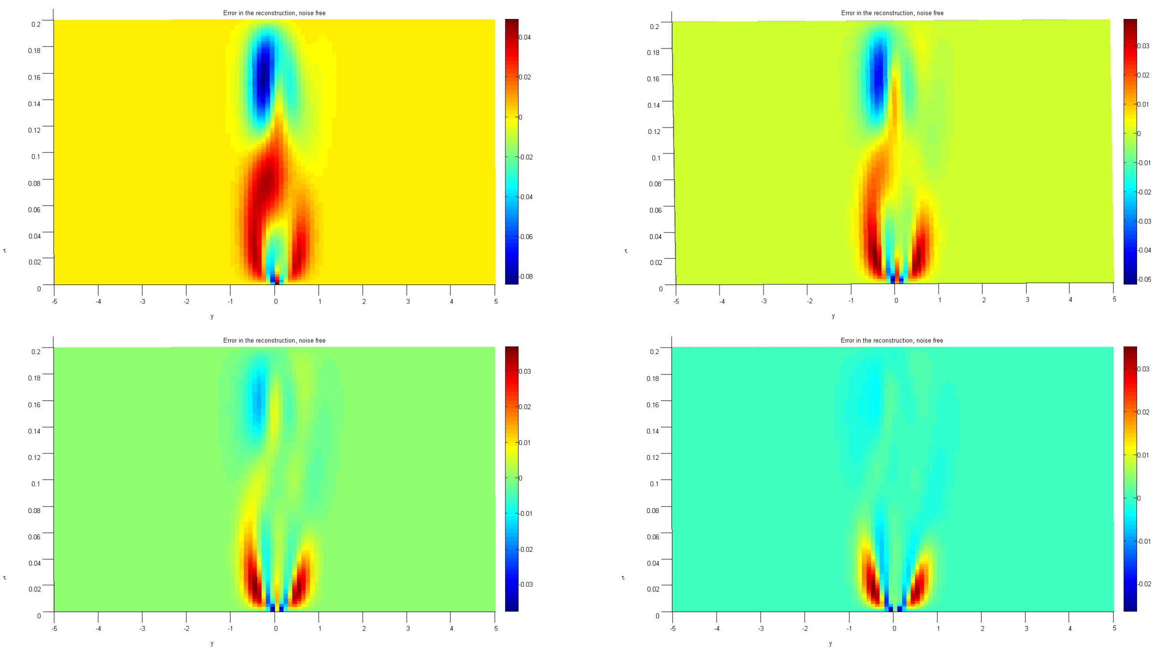

Figure 3.5: The error in the reconstruction after 500, 1000, 2000 and 4000 iteration steps.

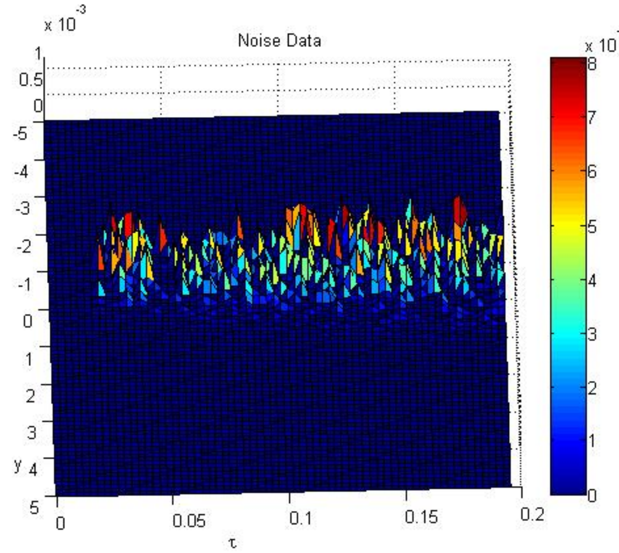

Figure 3.6: Noise data of 0.5% concentrated in a region of the form $y \in [-5/3, 5/3]$ and $\tau \in [1/30, 1/5]$.

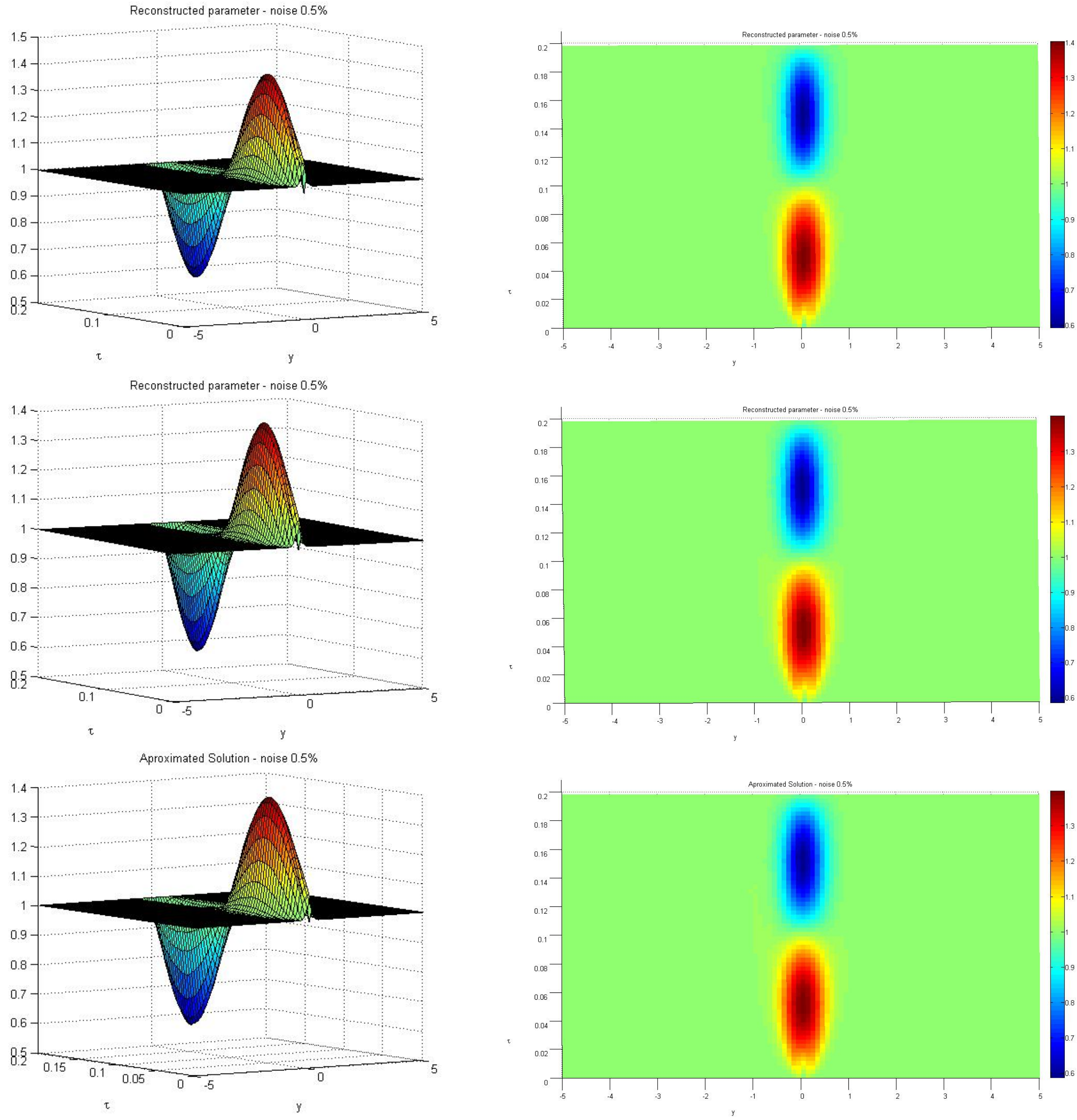

Figure 3.7: Reconstructed parameter after 800, 1200 and at the stopping iterate. Noise level of 0.5% concentrated in a region of the form $y \in [-5/3, 5/3]$ and $\tau \in [1/30, 1/5]$ concentrated near $y = 0$. On the left, the surface plot of the reconstructed solution. On the right, the intensity plot.

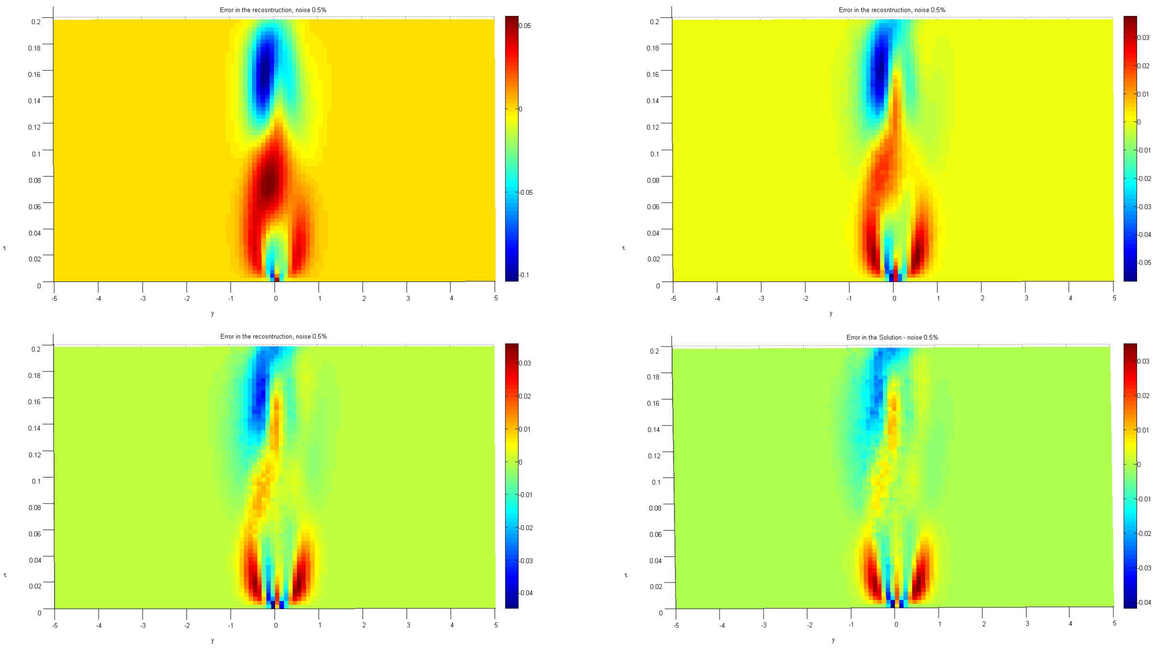

Figure 3.8: Error in the reconstruction after 400, 800, 1200 and at the stopping iterate.

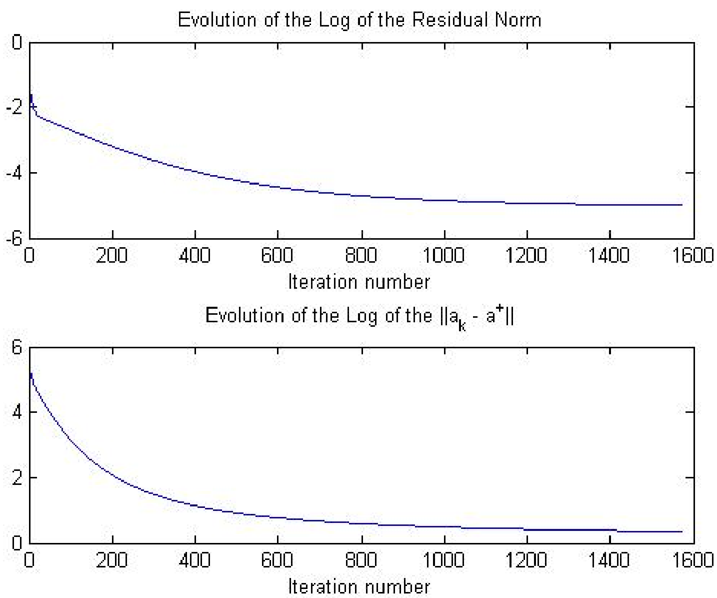

Figure 3.9: Evolution of the residual and the error in the approximated $L^2(\Omega)$-norm.

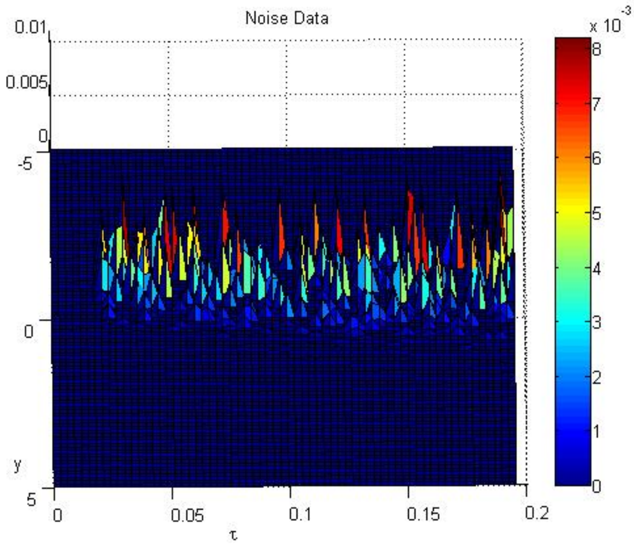

Figure 3.10: Noise data of 5% concentrated in a region of the form $y \in [-5/3, 5/3]$ and $\tau \in [1/30, 1/5]$.

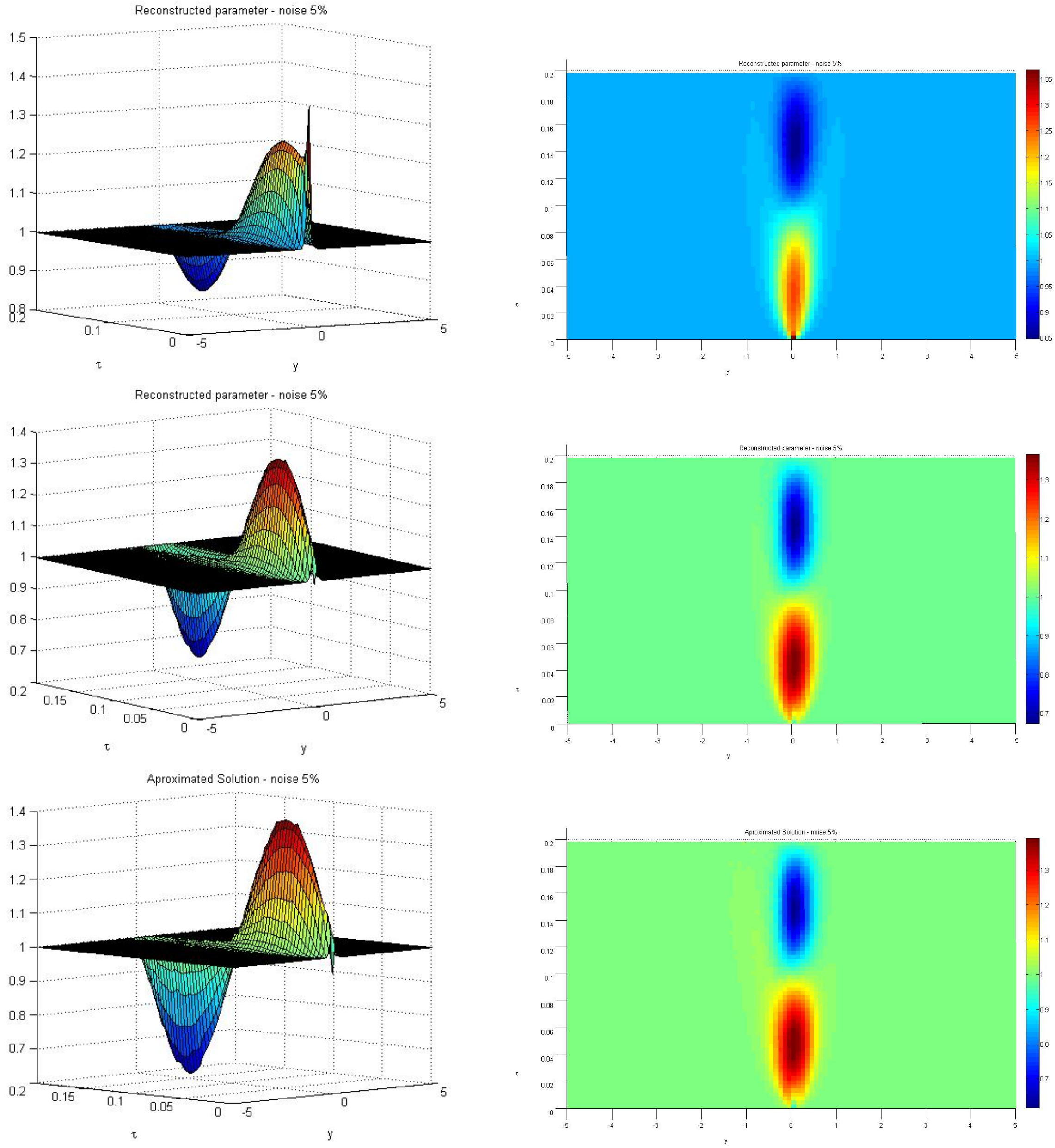

Figure 3.11: Reconstructed parameter after 100, 300 and at stopping final iteration. On the left, the reconstructed solution as a surface. On the right, the intensity plot.

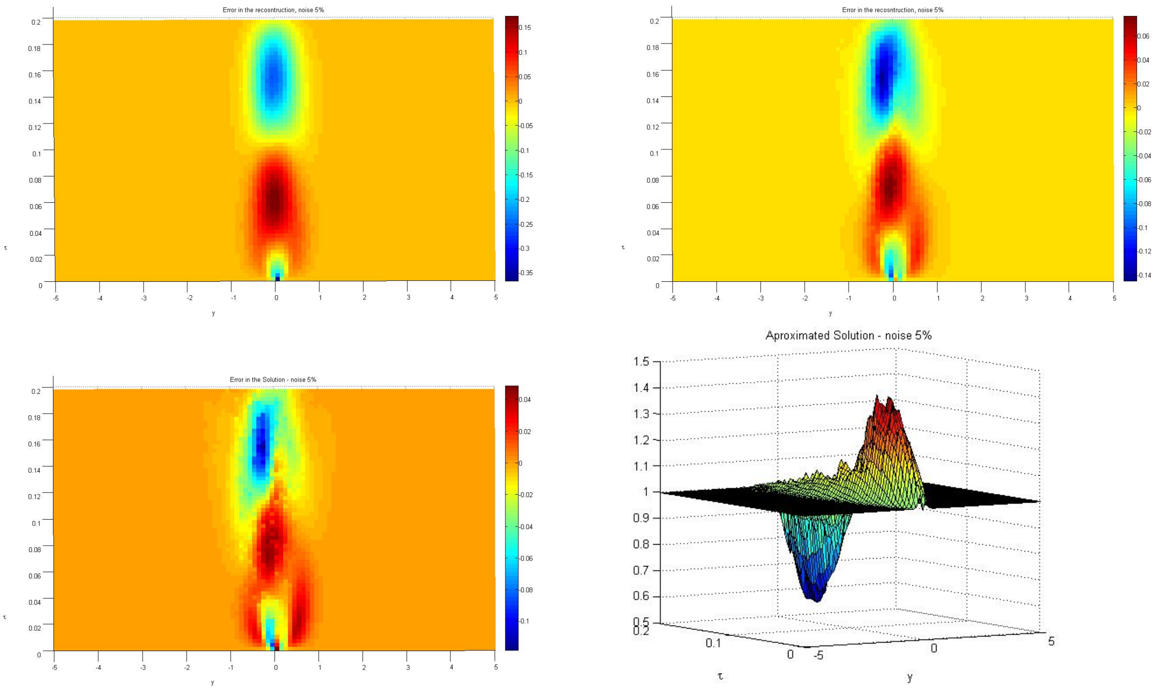

Figure 3.12: Error in the reconstruction for the respective iterations given in Figure 3.11.

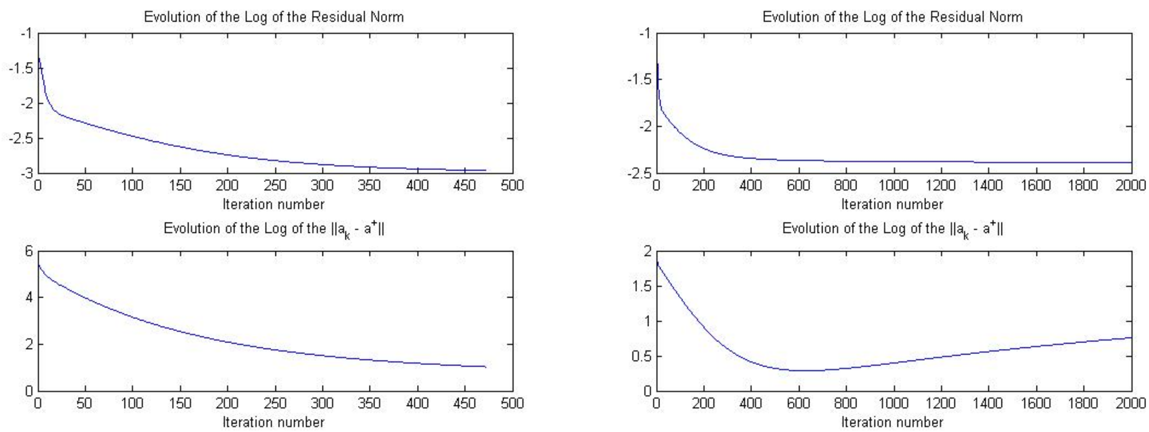

Figure 3.13: Evolution of the residual and the error in the approximate $L^2(\Omega)$-norm. In the left we can see the stopping index. In the right we can see the monotonicity properties of the iterates.

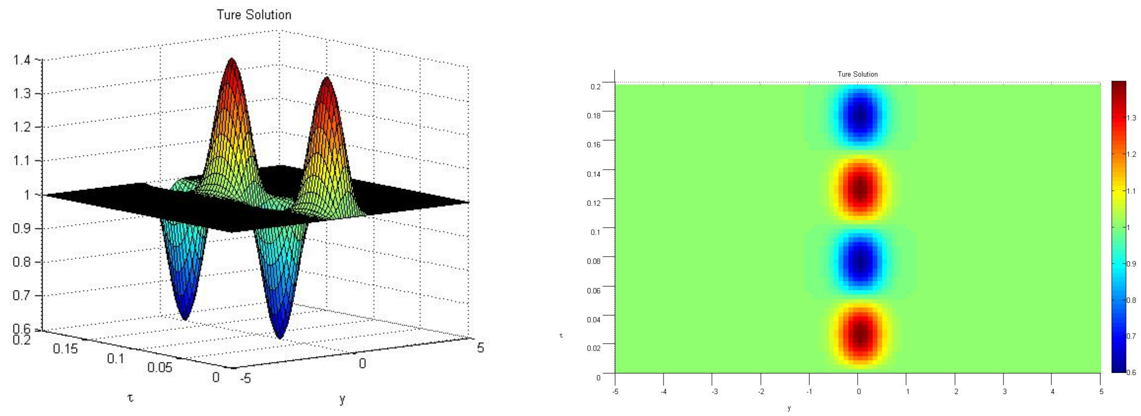

Figure 3.14: True parameter. On the left, as a surface, on the right as an intensity.

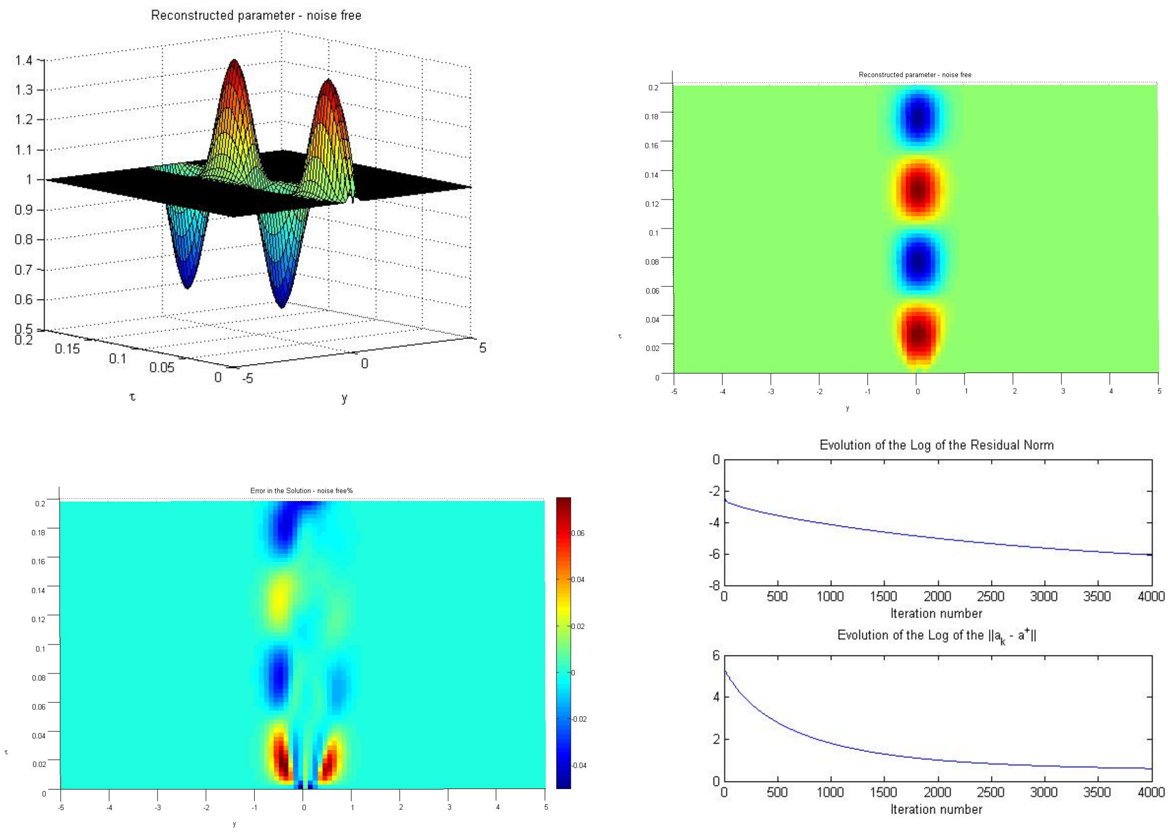

Figure 3.15: On the top, the reconstructed solution after 4000 iteration steps. On the bottom left, the error in the reconstruction. On the bottom right, the evolution of the residual and the error in the approximate $L^2(\Omega)$-norm.

Chapter 4

Relationship with Convex Risk Measures

In this chapter, we show a possible connection between convex risk measures and the interpretation of source condition (2.7). The main point is that we present a construction that allows us to associate the convex regularization functional f involved in the source condition to a convex risk measure. This circle of ideas is novel, to the best of our knowledge, and deserves careful further investigations.

From a heuristic perspective, a financial interpretation of the source condition (2.7) is carried out so that we have a restriction that allows us to quantify the risk associated to a given volatility level. By this we mean that upon computing the corresponding Black-Scholes solution as a function of the volatility, we are quantifying how much risk one has in the space of random variables associated to such volatility. In other words, we prove the following:

Theorem 4.0.2. *The source condition* (2.7) *can be interpreted as an a priori assumption on the risk associated to the correspondent position, given the volatility level.*

Proof. The proof is given at the end of Section 4.2. □

In financial practice, a number of ways have been proposed to assess the risk of a given portfolio or investment choice [71]. Perhaps the most well-known is the so-called *value at risk* (VaR), which is defined as follows: For a given portfolio, probability level and time period, the (VaR) is defined as the threshold value such that the probability of loss on the portfolio over the given time period exceeds this value is the given probability level. In other words, given a confidence level $\alpha \in (0, 1)$, the (VaR) of the portfolio at the confidence level α is given by the smallest number l such that the probability that the loss L exceeds l is not larger than $(1-\alpha)$

$$\mathrm{VaR}_\alpha = \inf\{l \in \mathbb{R} \,:\, P(L > l) \leq 1 - \alpha\}\,.$$

After some consideration it seems clear that the higher the (VaR) the higher the risk, and, in principle, the more undesirable such an investment would be. It turns out that the (VaR) has a serious pitfall, namely, it does not encourage diversification. This is related to the fact that it is not in general a convex function of the portfolio choice.

Several authors have developed theories about desirable properties for risk measures. See [71] and references therein. One of the most popular is the concept of a convex risk measure. It represents a quantitative assessment of the risk involved in the investor's preference for a financial position. Usually a position is described by the resulting discounted net worth at the end of a given period. Thus, it is represented by a random variable ν in a suitable probability space Γ. More precisely, we denote by $\mathcal{X} := \{\nu\, \Gamma \longrightarrow \mathbb{R}\}$ a convex set of real-valued random variables over all possible scenarios. Following [6, 42, 43, 44] we shall now introduce the definition of convex risk measure and give a brief explanation of its meaning.

Definition 5. *A map $\rho : \mathcal{X} \longrightarrow \mathbb{R}$ will be called a convex measure of risk if it satisfies the following conditions:*

- *Convexity.*
- *Non-increasing monotonicity, i.e., if the random variable ν_2 is dominated by the random variable ν_1 a.e., then $\rho(\nu_2) \geq \rho(\nu_1)$.*
- *Translation invariance, i.e., if $m \in \mathbb{R}$ is a deterministic variable in the sense that it takes the value m a.e., then*

$$\rho(\nu + m) = \rho(\nu) - m\,. \tag{4.1}$$

We now digress to give an intuitive interpretation of the different requirements above. The condition of convexity is related to risk aversion and it is important in diversifying risk. See [71] for details. The translation invariance condition is natural, since adding a deterministic quantity to a portfolio will decrease its risk in proportion to that quantity. The monotonicity says that if two portfolios ν_1 and ν_2 are such that for almost all events the return of ν_1 is greater than, or equal to, the return of ν_2, then the risk associated to ν_1 is smaller than the corresponding risk associated to ν_2.

4.1 Preliminary Results

In this section, we present the assumptions and preliminary results used later. The first assumption is that Ω is a bounded set. This is the same to assuming that the strikes K are bounded below and above by some positive constants. Moreover, we define the functional $f(a) = +\infty$ if $a \notin \mathcal{D}(F)$. Using the assumption of the existence of a source function $w^\dagger \in L_2(\Omega)$ that satisfies (2.7) and the definition of $\partial f(a^\dagger)$ we have that

$$f(a) - \langle w^\dagger, F'(a^\dagger)a\rangle \geq f(a^\dagger) - \langle w^\dagger, F'(a^\dagger)a^\dagger\rangle\,, \qquad \forall a \in U \text{ and } \forall w^\dagger \text{ s.t. } F'(a^\dagger)^* w^\dagger \in \partial f(a^\dagger)\,. \tag{4.2}$$

Let us set $g(-F'(a^\dagger)a) := \langle w, -F'(a^\dagger)a\rangle$. The existence of $w^\dagger$ satisfying (4.2) implies that it is the Lagrangian multiplier of

$$L : \mathcal{D}(F) \times L_2(\Omega) \longrightarrow \mathbb{R}$$
$$(a, w) \to f(a) + g(-F'(a^\dagger)a)\,,$$

i.e., it satisfies

$$L(a^\dagger, w) \leq L(a^\dagger, w^\dagger) \leq L(a, w^\dagger)\ .$$

However, it is not clear whether we have more than one $w^\dagger \in \mathcal{R}(F'(a^\dagger))$ satisfying (4.2). Indeed, it depends on the choice of f. For example, if f is differentiable on $a^\dagger$, then $\partial f(a^\dagger)$ is a single element. Then, from Lemma 1.4.2 it follows that $w^\dagger$ satisfies Equation (2.7) and therefore it is unique.

We define a family of separately convex functions (meaning that for a fixed w it is convex in a and vice versa) by

$$L_2(\Omega) \ni w \longmapsto h_w : \mathcal{D}(F) \longrightarrow \mathbb{R} \cup \{+\infty\}$$
$$a \longmapsto L(a, w) = f(a) + g(-F'(a^\dagger)a)\,. \tag{4.3}$$

Observe that $h_w(a)$ is a family of functions of the variable a depending on the parameter w.

Remark 4.1.1. *A particular property of* $h_{w^\dagger}$ *is that*

$$h_{w^\dagger}(a) - h_{w^\dagger}(a^\dagger) = L(a, w^\dagger) - L(a^\dagger, w^\dagger) = D_{\zeta^\dagger}(a, a^\dagger)\,.$$

However, this property holds true only in the special case when $w^\dagger$ *satisfies* (4.2)*.*

Remark 4.1.2. *Note that the source condition* (2.7) *together with the existence of an* f*-minimum norm solution for* (15) *is equivalent to the Karush-Kuhn-Tucker condition in convex optimization [38].*

Now, from the Fenchel conjugation theory [77, 85] we obtain a unique Fenchel conjugate function of h_w given by

$$\begin{aligned} \hat{h}_w^* :& L_2(\Omega) \longrightarrow \mathbb{R} \\ & v \longmapsto g^*(v) + f^*(-F'(a^\dagger)^* v)\,. \end{aligned} \tag{4.4}$$

If it happens that

$$g^*(v) = \begin{cases} 0 \text{ if } v = w \\ +\infty \text{ otherwise}\,. \end{cases}$$

Otherwise we would have difficulties in the above definition of $\hat{h}_w^*$. Hence, we focus on the related function h_w^* defined as

$$\begin{aligned} h_w^* : \mathbb{X} \subset L_2(\Omega) &\longrightarrow \mathbb{R} \\ v &\longmapsto h_w^*(v) := f^*(-F'(a^\dagger)^* v)\,, \end{aligned} \tag{4.5}$$

where $\mathbb{X} := \{v \in L^2(\Omega) \,:\, f^*(-F'(a^\dagger)^* v) \text{ is finite}\}$.

We note that since $\{0\} = \mathcal{N}(F'(a^\dagger)^*)$, then $h_w^*(0) = f^*(0) = 0$.

Lemma 4.1.1. *The functional* h_w^* *satisfies the convexity and monotonicity axioms.*

Proof. The convexity follows directly from the properties of the Fenchel conjugate function [85, Theorem 2.3.1]. To prove the monotonicity: let $v_1, v_2 \in \mathbb{X}$ satisfy $v_1 \geq v_2$. From the definition of the Fenchel conjugate we have $h_w^*(v) = f^*(-F'(a^\dagger)^* v) \geq \langle a, -F'(a^\dagger)^* v\rangle - f(a)$. Positivity of $F'(a^\dagger)a$ (see [22, Theorem 4.2]) implies that

$$\begin{aligned} 0 \leq \langle F'(a^\dagger)a, v_1 - v_2\rangle &= \langle F'(a^\dagger)a, v_1\rangle + f(a) - (\langle F'(a^\dagger)a, v_2\rangle + f(a)) \\ &\leq -h_w^*(v_1) + h_w^*(v_2)\,. \end{aligned}$$

□

In the sequel we give a construction of a convex risk measure ρ in the present context. This will be achieved using the properties of h_w^* and an interesting probabilistic representation of $v \in \mathbb{X}$ coming from Malliavin Calculus [45].

We start by associating our notation with that of [45]. Equation (9) is associated to the diffusion process $\{y_t \,:\, 0 \le t \le T\}$ that satisfies the dynamics

$$dy_t = \left(r - q - \frac{\sigma(t,y_t)^2}{2}\right) dt + \sigma(t,y_t) dW_t \,, \qquad y_{t_0} = y_0 \,, \tag{4.6}$$

in the risk neutral probability measure $\mathbb{Q}$.

We recall that the process (4.6) is the diffusion (1) in logarithmic variables where $\sigma \longmapsto a \in \mathcal{D}(F)$ by (8).

For the sake of simplicity, we assume that the process (4.6) has no dividend nor interest rates, i.e., $b = 0$.

Following [45], denoted by $\{Y_t \,:\, 0 \le t \le T\}$ the first variation process associated to $\{y_t \,:\, 0 \le t \le T\}$ and defined by the stochastic differential equation

$$dY_t = (\sigma^2(Y_t))' Y_t dt + \sigma'(Y_t) dW_t \qquad Y_{t_0} = 1 \,.$$

Remark 4.1.3. *We now identify $\sigma^\dagger \longmapsto \sqrt{2a^\dagger}$ and $\tilde\sigma \longmapsto \sqrt{2\tilde a}$ given by (8) with $a^\dagger, \tilde a \in \mathcal{D}(F)$. Then, for sufficiently small $\varepsilon > 0$, the diffusion coefficient $\sigma^\dagger + \varepsilon\tilde\sigma$ satisfies the uniform ellipticity condition*

$$\exists \eta > 0 \;:\; \zeta^T (\sigma^\dagger + \varepsilon\tilde\sigma)^T(x)(\sigma^\dagger + \varepsilon\tilde\sigma)(x)\zeta \ge \eta|\zeta|^2,$$

for all $\zeta \in \mathbb{R}^2$ and for all $x \in \Omega$.

We introduce the auxiliary set

$$\Gamma := \left\{ \Theta \in L^2[0,T] \,|\, \int_0^T \Theta(t) dt = 1 \right\} ,$$

which contains for example the constant function $\Theta(t) = 1/T$.

Our first result is a representation lemma.

Lemma 4.1.2. *Let $v \in \mathcal{R}(F'(a^\dagger))$. Then, there exists a random variable $\pi_{a^\dagger}$ such that*

$$v = \mathbb{E}_{\mathbb{Q}}^{y_0}[\Phi(y_t)\pi_{a^\dagger}], \tag{4.7}$$

where $\mathbb{Q}$ is the risk neutral probability measure.

Proof. Let

$$\tilde{\beta}_\Theta = \Theta(t)(\beta(T) - \beta(0))\chi_{0\le t\le T}$$

where $\{\beta(t) : 0 \le t \le T\}$ is the process given in [45, Lemma 3.1].

Since $\sigma^\dagger + \varepsilon\tilde{\sigma}$ satisfies the uniform ellipticity condition (see Remark 4.1.3) we have from [45, Proposition 3.3] that the Gâteaux derivative at $\sigma^\dagger$ in the direction $\tilde{\sigma}$ is given by

$$\mathbb{E}_{\mathbb{Q}}^{y_0}[\Phi(y_t)D_t^*((\sigma^\dagger)^{-1}(y_t)Y_t\tilde{\beta}_\Theta(T))],$$

where $D_t^*((\sigma^\dagger)^{-1}(y_t)Y_t\tilde{\beta}_\Theta(T))$ is the Skorohod integral [73] of the possibly anticipative process

$$\{(\sigma^\dagger)^{-1}(y_t)Y_t\tilde{\beta}_\Theta(T) : 0 \le t \le T\}$$

for any $\Theta \in \Gamma$. □

We remark that the linearity of D_t^* with respect to $\tilde{\sigma}$ arises through the process β_t. See Proposition 3.3 of [45].

Lemma 4.1.3. *The constants do not belong to $\mathcal{R}(F'(a^\dagger))$.*

Proof. If $1 \in \mathcal{R}(F'(a^\dagger))$, then there exists $h \in \mathcal{D}(F'(a^\dagger))$ such that $F'(a^\dagger)h = 1$. Thus, 1 would satisfy (1.14), i.e.,

$$0 = 1_\tau + a^\dagger(1_{yy} - 1_y) = h(u_{yy} - u_y).$$

Using the same argument in the proof of Lemma 1.4.1 we have that $(u_{yy} - u_y)$ cannot vanish in a set of positive measures. Thus $h = 0$ a.e. This is a contradiction with the fact that $F'(a^\dagger)h = 1$ since $F'(a^\dagger)$ is linear. □

At this point, we have two interesting sets of random variables for our convex risk measure construction. Firstly,

$$\mathcal{X} := \{\nu + m : \nu = \Phi(y_t) \text{ and } m \in \mathcal{C}\}$$

and secondly,

$$\mathcal{X}_1 := \{\pi_{a^\dagger} + m \, : \, \pi_{a^\dagger} = D_t^*((\sigma^\dagger)^{-1}(y_t)Y_t\tilde{\beta}_\Theta(T)) \text{ and } m \in \mathcal{C}\} \, ,$$

where $\mathcal{C}$ is the set of all constants.

Remark 4.1.4. *It follows from Lemma 4.1.2 that we have a representation of* $\mathbb{X}$ *by* $\mathcal{X}$ *and* $\mathcal{X}_1$ *given by the weighted expectation* $\mathbb{E}_{\mathbb{Q}}^{y_0}[\cdot]$ *with weight* $D_t^*((\sigma^\dagger)^{-1}(y_t)Y_t\tilde{\beta}_\Theta(T))$ *and* $\Phi(y_t)$ *respectively. We remark that the terminology weight here is used in a loose sense, since indeed* $D_t^*((\sigma^\dagger)^{-1}(y_t)Y_t\tilde{\beta}_\Theta(T))$ *may have a negative value.*

The following lemma plays a central role in our analysis below.

Lemma 4.1.4. *If* $\nu \equiv 1$, *then*

$$\mathbb{E}_{\mathbb{Q}}^{y_0}[\nu D_t^*((\sigma^\dagger)^{-1}(y_t)Y_t\tilde{\beta}_\Theta(T))] = 0 \, .$$

Proof. This follows directly from the duality between the Skorohod integral and the Malliavin derivative [73], and the fact that $D_t 1 = 0$. □

4.2 Convex Risk Measures

We are now ready to state the main results from this chapter.

Proposition 4.2.1. *[First alternative for a convex risk measure] The functional*

$$\rho \, : \mathcal{X} \longrightarrow \mathbb{R} \qquad \nu \longmapsto \rho(\nu) := h_w^*(\mathbb{E}_{\mathbb{Q}}^{y_0}[\nu \cdot \pi_{a^\dagger}]) - \mathbb{E}_{\mathbb{Q}}^{y_0}[\nu] \tag{4.8}$$

satisfies the convex risk measure axioms.

Proof. From the linearity of the expectation operator and the properties of the functional h_w^* in Lemma 4.1.1, the convexity and monotonicity axioms follow.

In order to prove the translation axiom, we write

$$\tilde{\rho} \, : \mathcal{X} \longrightarrow \mathbb{R} \qquad \nu \longmapsto \tilde{\rho}(\nu) := h_w^*(\mathbb{E}_{\mathbb{Q}}^{y_0}[(\nu - \mathbb{E}_{\mathbb{Q}}^{y_0}[\nu]) \cdot \pi_{a^\dagger}]) - \mathbb{E}_{\mathbb{Q}}^{y_0}[\nu] \, .$$

Let $\nu + m \in \mathcal{X}$. From the linearity of the expected value

$$\begin{aligned}\tilde{\rho}(\nu+m) &= h_w^*(\mathbb{E}_{\mathbb{Q}}^{y_0}[(\nu+m-\mathbb{E}_{\mathbb{Q}}^{y_0}[\nu+m])\cdot\pi_{a^\dagger}]) - \mathbb{E}_{\mathbb{Q}}^{y_0}[\nu+m] \\ &= h_w^*(\mathbb{E}_{\mathbb{Q}}^{y_0}[(\nu-\mathbb{E}_{\mathbb{Q}}^{y_0}[\nu])\cdot\pi_{a^\dagger}]) - \mathbb{E}_{\mathbb{Q}}^{y_0}[\nu] - m = \tilde{\rho}(\nu) - m\,.\end{aligned}$$

Hence $\tilde{\rho}$ satisfies the translation axiom.

Now we show that $\tilde{\rho} = \rho$. Indeed, by definition, $\mathcal{X} = \mathcal{D}(\tilde{\rho}) = \mathcal{D}(\rho)$. Let us now take $\nu \in \mathcal{X}$; by definition of expectation $\mathbb{E}_{\mathbb{Q}}^{y_0}[\nu] = c$ where c is a constant, it follows from Lemma 4.1.4 that

$$\begin{aligned}\tilde{\rho}(\nu) &= h_w^*(\mathbb{E}_{\mathbb{Q}}^{y_0}[(\nu-\mathbb{E}_{\mathbb{Q}}^{y_0}[\nu])\cdot\pi_{a^\dagger}]) - \mathbb{E}_{\mathbb{Q}}^{y_0}[\nu] \\ &= h_w^*(\mathbb{E}_{\mathbb{Q}}^{y_0}[\nu\cdot\pi_{a^\dagger}] - \mathbb{E}_{\mathbb{Q}}^{y_0}[c\cdot\pi_{a^\dagger}]) - \mathbb{E}_{\mathbb{Q}}^{y_0}[\nu] = \rho(\nu) \quad \text{for all} \quad \nu\in\mathcal{X}\,.\end{aligned}$$

Thus $\tilde{\rho} = \rho$. □

Proposition 4.2.2. *[Second alternative for a convex measure of risk] The functional*

$$\rho_1 : \mathcal{X}_1 \longrightarrow \mathbb{R} \qquad\qquad \pi \longmapsto \rho_1(\pi) := h_w^*(\mathbb{E}_{\mathbb{Q}}^{y_0}[\nu\cdot\pi])\,, \tag{4.9}$$

satisfies the convex risk measure axioms.

Proof. Using the same argument of Proposition 4.2.1, the convexity and monotonicity axioms follow.

In order to prove the translation axiom, we write

$$\tilde{\rho}_1 : \mathcal{X}_1 \longrightarrow \mathbb{R} \qquad\qquad \pi \longmapsto \tilde{\rho}_1(\pi) := h_w^*(\mathbb{E}_{\mathbb{Q}}^{y_0}[\nu\cdot(\pi - \mathbb{E}_{\mathbb{Q}}^{y_0}[\pi])]) - \mathbb{E}_{\mathbb{Q}}^{y_0}[\pi]\,.$$

Then, for $\pi + m \in \mathcal{X}_1$, from the linearity of the expectation operator we have that

$$\begin{aligned}\tilde{\rho}_1(\pi+m) &= h_w^*(\mathbb{E}_{\mathbb{Q}}^{y_0}[\nu\cdot(\pi+m-\mathbb{E}_{\mathbb{Q}}^{y_0}[\pi+m])]) - \mathbb{E}_{\mathbb{Q}}^{y_0}[\pi+m] \\ &= h_w^*(\mathbb{E}_{\mathbb{Q}}^{y_0}[\nu\cdot(\pi-\mathbb{E}_{\mathbb{Q}}^{y_0}[\pi])]) - \mathbb{E}_{\mathbb{Q}}^{y_0}[\pi] - m = \tilde{\rho}_1(\pi) - m\,.\end{aligned}$$

Hence, $\tilde{\rho}_1$ satisfies the translation axiom.

By definition, $\mathcal{X}_1 = \mathcal{D}(\tilde{\rho}_1) = \mathcal{D}(\rho_1)$. Let us take $\pi \in \mathcal{X}_1$. From Lemma 4.1.4 we conclude that $\mathbb{E}_{\mathbb{Q}}^{y_0}[\pi] = \mathbb{E}_{\mathbb{Q}}^{y_0}[1\cdot\pi] = 0$.

Thus, $\tilde{\rho}_1(\pi) = \rho(\pi) \quad \text{for all} \quad \pi \in \mathcal{X}_1\,.$

□

We note that the choice of $\sigma^\dagger$ enters in a crucial and nonlinear way in the convex risk measure. Furthermore, the source condition (2.7) allows us to construct convex risk measures in the spaces of random variables associated to the diffusion process (4.6).

Proof of Theorem 4.0.2. Given (2.7) satisfied, Propositions 4.2.1 and 4.2.2 show the existence of a convex risk measure for the associated position. Therefore, the assertion of Theorem 4.0.2 follows. □

4.2.1 A Convex Risk Measure Associated with the Boltzmann-Shannon Entropy

We now illustrate the construction of the convex risk measure by considering the process (4.6) under constant volatility with vanishing interest and dividend rates. For this particular case, the representation (4.7) (or the *vega* in financial terms) is given by the formula (see [45])

$$
\begin{aligned}
&\mathbb{E}_{\mathbb{Q}}^{y_0}\left[\Phi\left(y\exp\left(\sigma^\dagger W_\tau - \frac{(\sigma^\dagger)^2}{2}\tau\right)\right)\cdot\left(\frac{W_\tau^2}{\sigma^\dagger\tau} - W_\tau - \frac{1}{\sigma^\dagger}\right)\right] \\
&\quad = \int_\Omega dz\,d\tau\; p(z,\tau)\Phi\left(y\exp\left(\sigma^\dagger z - \frac{(\sigma^\dagger)^2}{2}\tau\right)\right)\cdot\left(\frac{z^2}{\sigma^\dagger\tau} - z - \frac{1}{\sigma^\dagger}\right), \qquad (4.10)
\end{aligned}
$$

where $p(z,\tau) = e^{-\frac{z^2}{2\tau}}/\sqrt{2\pi\tau}$ is the Gaussian probability density function.

Let us take $v \in \mathbb{X}$ and compute $F'(a^\dagger)^* v$. By Fubini's Theorem

$$
\begin{aligned}
&\langle F'(a^\dagger)a, v\rangle \\
&= \int_\Omega d\tau'\,dy v(\tau',y)\int_\Omega d\tau\,dz p(z,\tau)\Phi\left(y\exp\left(\sigma^\dagger z - \frac{(\sigma^\dagger)^2}{2}\tau\right)\right)\cdot\left(\frac{z^2}{\sigma^\dagger\tau} - z - \frac{1}{\sigma^\dagger}\right) \\
&= \int_\Omega d\tau\,dz p(z,\tau)\left(\frac{z^2}{\sigma^\dagger\tau} - z - \frac{1}{\sigma^\dagger}\right)\int_\Omega d\tau'\,dy v(\tau',y)\Phi\left(y\exp\left(\sigma^\dagger z - \frac{(\sigma^\dagger)^2}{2}\tau\right)\right)
\end{aligned}
$$

Thus,

$$
-F'(a^\dagger)^* v = \left(\frac{z^2}{\sigma^\dagger\tau} - z - \frac{1}{\sigma^\dagger}\right)\langle -v, \Phi(\cdot)\rangle. \qquad (4.11)
$$

We now consider the regularization functional f as the Boltzmann-Shannon entropy

$$f(a) = \int_\Omega a \log(a)\, dx\,, \qquad a \in \mathcal{D}(F)\,,$$

whose Fenchel conjugate is given by

$$f^*(\mu) = \int_\Omega e^{\mu-1}\, d\tilde{x}\,.$$

Since we are in a Gaussian model, applying [2, Lemma 11] and (4.11) to the definition of ρ with $\nu = \Phi\left(y \exp\left(\sigma^\dagger(z) - (\sigma^\dagger)^2\tau/2\right)\right)$ we get

$$\rho(\nu) = -\log\left(\mathbb{E}^{y_0}_{\mathbb{Q}}\left[\exp\left(\frac{z^2}{\sigma^\dagger\tau} - z - \frac{1}{\sigma^\dagger}\right)\langle -v, \nu\rangle\right]\right) - \mathbb{E}^{y_0}_{\mathbb{Q}}[\nu]\,. \tag{4.12}$$

Compare with [43, Example 12].

Chapter 5

Conclusions and Future Directions

In this chapter, we draw some conclusions concerning the results developed in this work. We also discuss some future directions of work in this area.

In [22, 35], the Tikhonov regularization theory was developed to calibrate the local volatility that belongs to a more general admissible set. Indeed, they look for $a \in \mathcal{D}(F) \subset H^1(\Omega)$. One of the main difficulties to work with the admissible parameter class given by (11) is that, $\mathcal{D}(F)$ has no interior points in the $H^1(\Omega)$-norm. This is due to the fact that we are in the critical Sobolev embedding exponent for $\Omega \subset \mathbb{R}^2$. See [1].

Within this framework, many results presented in this work remain true. On the other hand, some main results become unproved. More specifically, we have the following:

- Existence, uniqueness and the regularity estimates of the solution to the PDE equation (9) with initial condition (10) developed at Section 1.2 remain true by use similar arguments. See also the Appendix in [35] or reference [22].

- Theorem 1.3.1 of Section 1.3 also holds. Moreover, there exists a directional derivative $F'(a) \cdot h$ in all directions $h \in H^1(\Omega)$ such that $a + h \in \mathcal{D}(F) \subset H^1(\Omega)$, as the Lemma 1.3.1 shows. Since, $\mathcal{D}(F)$ has no interior points when equipped with the $H^1(\Omega)$-norm, $F'(a)$ is not necessarily differentiable in any direction $h \in H^1(\Omega)$. In other words, F is not Gateaux differentiable. This will not affect the convergence analysis of the Tikhonov regularization developed early. In fact, for such analysis we only need that the operator F admits a one-side directional derivative at $a^\dagger$ in the directions $a - a^\dagger$, for all $a \in \mathcal{D}(F)$. See Subsection 2.1.1 at Chapter 2. The sufficient condition in this case is that, for all $a \in \mathcal{D}$, there exists $t_0 > 0$ such that

$$a^\dagger + t(a - a^\dagger) = ta + (1-t)a^\dagger \in \mathcal{D}(F) \qquad \forall t \leq t_0 \ .$$

As $\mathcal{D}(F)$ is convex, the requirement above is satisfied. Thanks to the convexity of $\mathcal{D}(F)$, the operator $F'(a^\dagger)$ has properties that mimic the Gâteaux derivative. In particular, there is an adjoint operator $F'(a^\dagger)^* : (W_p^{1,2}(\Omega))^* \longrightarrow \mathcal{H}(\Omega)$ defined as

$$\langle F'(a^\dagger)^* v, a\rangle = \langle v, F'(a^\dagger)a\rangle_{((W_p^{1,2}(\Omega))^*, W_p^{1,2}(\Omega))}, \quad a \in \mathcal{H}(\Omega), \quad v \in (W_p^{1,2}(\Omega))^* .$$

- The loss of the Fréchet derivative properties of F affect the convergence rates of the Tikhonov regularization. Indeed, without the assumption of a continuous Fréchet differentiability of the operator F, we are not able to prove Lemmas 1.4.1 and 1.4.2. Hence, we do not have Lemma 2.1.1. Moreover, we are not able to prove the main result on iterative methods, the tangential cone condition properties (see Theorem 1.4.2). Without continuity of the Fréchet derivative of F we are not able to prove the convergence of the Landweber iteration in Chapter 3.

In Chapter 1, we analyzed the existence and uniqueness of solutions to the PDE (9) with initial condition (10). Also, we verified well known properties of the parameter-to-solution map F with parameters in the admissible class $\mathcal{D}(F) \subset \mathcal{H}(\Omega)$. These properties imply the ill-posedness of the calibration problem. Moreover, we characterized the sets $\mathcal{N}(F'(a^\dagger))$ and $\mathcal{R}(F'(a^\dagger)^*)$ as $L^2(\Omega)$ subsets. This characterization implies in the existence of an approximate source condition to the problem and, consequently, in convergence rates. The main novelty is the proof of the *local tangential cone condition* for the parameter-to-solution map F. See Theorem 1.4.2. The local tangential cone condition (1.21) and the uniform limitation of $F'(\cdot)$ (see equation 1.16) are enough to prove regularization properties of iterative methods during the Chapter 3.

In Chapter 2, we prove existence and convergence results for the regularized solutions of the inverse problem associated to the calibration of local volatility surface in European option prices. This is done using Tikhonov regularization. The main novelty was the use of a regularization term that only requires convexity properties and weak lower-semicontinuity. Thus, the present regularization applies to a large class of regularization functionals.

A general theoretical framework based on Tikhonov regularization by means of convex penalizing functions is an extension of the quadratic regularization that has been previously studied in the Inverse Problem literature [35]. We prove the existence of approximate source condition (2.11) for the regularization problem under consideration. In particular, if the regularization functional is $f(\cdot) = \|\cdot\|^2_{\mathcal{H}(\Omega)}$, then the source condition (2.7) coincides with the representation that remained an open problem in [22, 35]. On the theoretical side, the strength is that this yields better convergence rates with respect to the noise level in the measurements. Furthermore, it allows for convergence in spaces different from those in the quadratic regularization setting. In fact, in some cases, the convergence of certain convex regularization expressions implies convergence in the L^1-norm.

Related to the regularization method under consideration, we establish for Bregman distances better convergence rates than those available in the literature to this problem. Another advantage of the current approach is the requirement of weaker conditions than those previously required in the literature. Namely, we only require (2.12). This analysis also allows us to obtain convergence of the regularized solution with respect to the noise level in $L^1(\Omega)$ by means of a Kullback-Leibler regularization functional. See Equation (2.20). The Bregman distance regularization is motivated in Section 2.2 using exponential families. The construction of the Tikhonov functional (2.15) is heuristic. On the other hand, the convergence analysis of the regularized solutions is completely rigorous.

The same convergence results hold true if we measure the misfit of the Tikhonov functional (2.1) in $W_p^{1,2}(\Omega)$. The intuition behind the use of the $W_p^{1,2}(\Omega)$ norm is that we have continuous dependence of the Tikhonov functional with respect to information not only about the prices but also with respect to the sensitivities u_τ, u_{yy}, and u_y. Those are the so-called Greeks in financial engineering. On the other hand, we need more information on the measurement data u^δ.

In Chapter 3, we prove regularizing properties of approximated solutions to the calibration problem given by iterative methods. The main novelty is the use of an iterative regularization of the local volatility surface.

In the beginning we analyze the nonlinear Landweber iteration to the calibration problem. Originally, the Landweber iteration holds with the discrepancy principle measured in the $W_2^{1,2}(\Omega)$-norm and $F'(\cdot)^*$ evaluated in the $W_2^{1,2}(\Omega)$-inner product. Given the difficulties in calculating $F'(\cdot)^*$ in the $W_2^{1,2}(\Omega)$-inner product we recover the convergence analysis of the iteration in the $L^2(\Omega)$ - inner product, making use of continuous Sobolev embedding Theorems and a discrepancy principle in the $L^2(\Omega)$-norm. Moreover, in Section 3.4, we reformulate the inverse problem of calibrating the local volatility in a system of nonlinear ill-posed problems. Kaczmarz-type strategies to solve nonlinear ill-posed problems in the literature are revised. The possibility of applications to inverse problem under consideration is obtained.

A heuristic financial interpretation of the source condition (2.7) is that we have a restriction that allows us to quantify the risk associated to a given volatility level. By this we mean that upon computing the corresponding Black-Scholes solution as a function of the volatility, we are quantifying how much risk one has in the space of random variables associated to such volatility. This is done by means of the source condition (2.7). Indeed, in Chapter 4, we constructed a functional that through the Fenchel duality defines different convex risk measures. The availability of such risk measures permits quantifying the risk associated to random variables and portfolios of the underlying model.

In Apendice A, we do a brief review about no-arbitrage principles and its relations to volatility surface calibration proposed in this work. We motivate a shape restriction on the option price surfaces, using the Tikhonov regularization approach derived in Section 2.2. Moreover, using the penalization $f(a) = \|a - a_0\|^2_{\mathcal{H}(\Omega)}$ in the Tikhonov functional and properties of the calibration problem, we obtain better convergence rates to the regularized solutions than those obtained in [22, 35].

Some directions of future research would be:

- The numerical implementation of the present results with actual market data. An implementation for the case of the standard quadratic Tikhonov regularization can be found in [35, 59].
- The validation of the results obtained here to American options [65, 63]. This is a free boundary problem [46]. An interesting inverse problem related to American options is the identification of the free boundary interface. Related to this, are the well known level set approach developed in [26, 25, 27] and references therein.
- Verification of the *local tangential cone condition* to others diffusive parabolic problems [41].
- Explore further the connection to risk measures [42, 43]. In particular the freedom in the choice of such measures.

Appendix A

Tikhonov Regularization and No-Arbitrage Conditions

The development of mathematical methods for pricing derivatives was a major player in the expansion of derivative markets. The absence of arbitrage opportunities in an equilibrium market situation is a fundamental principle underlying the modern theory of financial asset pricing. We recall that an arbitrage portfolio consists of a portfolio whose value V_t at $t = 0$ vanishes, and such that $\mathbb{P}[V_T \geq 0] = 1$ and $\mathbb{P}[V_T \neq 0] > 0$ [63]. If the nature of uncertainty can be described by a stochastic process as (1), then the absence of arbitrage opportunities implies that there exists a risk-neutral density $p(S_T|S_t, T, t, r)$ such that the European call option price at time t has a probabilistic representation given by (2). See, for example, [63, 65].

Call option prices are, usually, quoted by their Black-Scholes implied volatility, i.e., the unique volatility parameter value for with the Black-Scholes formula yields the observable option price. Under the no-arbitrage principle, a European option price can be expressed as a function of five parameters: the current underlining price S_0, the time to expiration τ, the strike price K, the risk-free interesting rates r and the volatility parameter σ. As it is well known, σ is the unique free parameter, i.e., is the unique quantity in the Black-Scholes model that cannot be directly observed.

Given a call option contract and its market price, we can always invert the Black-Scholes formula to get the volatility parameter σ. This inversion is ill-posed in the sense of Hadamard. See Subsection 1.5. This option price model inverse problem has attracted intensive attention in the last several years [7, 18, 22, 28, 35, 55]. The interesty of the models proposed in the literature is to price European options correctly given a set of observations as in (12). A model which can price correctly all the relevant option or can capture the skewness seems to be able to give more accurate risk sensitivity calculations. The latter are the so called *Greeks* in mathematical finance [63, 65].

A.1 Motivation and some Definitions

Absence of arbitrage imposes shape restrictions on the surface of the option prices function. See, for example, [20, 62, 80, 82].

Following [20, 62, 80, 82], an arbitrage opportunity is a self-financing portfolio of securities that has a negative value today and a nonnegative value at a given time in the future independent of the market behavior, with positive probability.

In a model with implied volatility given for all maturities and strikes there is no-arbitrage in the input if, and only if, the following properties hold [20, 62, 80, 82]:

1. For a given maturity, the call price is non-increasing and convex with respect to the strike.

2. The call price is a decreasing function of time.

Assuming the underlying volatility to be constant in the Black-Scholes model of the European call option, it is well known that the price of the option satisfies (2). Taking the first derivative with respect to K in (2), it follows that

$$\frac{\partial U^{T,K}(S_t, K, r, \sigma^2)}{\partial K} = -\exp(-r\tau)\int_K^\infty \rho(S_T|S_t,\tau,r)dS_T\,, \tag{A.1}$$

where $\tau = T - t$. Differentiating the call price U twice with respect to K, we have

$$\frac{\partial^2 U^{T,K}(S_t, K, r, \sigma^2)}{\partial K^2} = \exp(-r\tau)\rho(K|S_t,\tau,r)\,. \tag{A.2}$$

Since the density function ρ is non-negative, it follows that the call price U must be non-increasing and convex with respect to K. The monotonicity of the price $U^{T,K}$ with respect to maturity T follows directly by the Black-Scholes formula [15]. Its means that, in a constant volatility model, the European option price is arbitrage-free, i.e., satisfies the Items 1 and 2 above.

The dependence of the implied surface volatility quoted on European options from maturity and strikes implies in a phenomena referred as *smile effect* [33, 34]. In the presence of smiles, arbitrage opportunities may exist among the quoted options, even if there is no-arbitrage in the original set of observed option prices. Because of this, the inverse models may use some algorithms to correct the arbitrage opportunities. See [5, 31, 62, 82] and references in there.

To make the things worse, in a typical option market, often, the prices are observed in very few different maturities and strike levels. Moreover, some of those option contracts are not liquid at all. As many of the option price inverse problem models require a complete set of European option price observations, over a continuous of maturities and strikes, a good method of interpolation for the option prices is required [22, 28, 35, 68, 82]. For a given discrete set of observations satisfying Items 1 and 2, there are many interpolation methods that preserve monotonicity and convexity properties.

Standard market practices include to interpolate the implied volatility, then substitute the implied volatility into the Black-Scholes equation to quote the option prices. Cubic-spline volatility interpolation is a common method used in option price models. See [62, 82] and references in there. In [82] a cubic B-spline interpolation method was proposed to preserve the shape of the option price function that minimizes the distance between the implied risk-neutral density and the prior approximation in L^2-norm. The authors of [82] assume that the volatility depends only on the strike, and the set of observations satisfies Items 1 and 2. Kahalé [62] developed an interpolation strategy to interpolate the local volatility surfaces that produces prices of options close to the input prices.

An interpolation strategy implies in introducing errors in the input data. As the volatility calibration problem is strongly sensitive to the small variations on the input data (see Lemma 1.5.1), some regularization methods need to be considered. On the other hand, measurement or numerical errors on the observation data can imply in a set of options that does not satisfy Items 1 and 2.

It is well known that for a model like (1) to be arbitrage-free, the coefficients $\nu(t, S(t))$ and $\sigma(t, S(t))$ cannot be arbitrarily specified, but they should be linked by certain relations. Those specifications are know as drift restrictions. Many works have reported characterizations of absence of arbitrage opportunities in terms of drift conditions. See [20, 80] and references in there.

As mentioned before, the calibration of the local volatility parameter in the European option price model has a fundamental importance. However, no-arbitrage principle implies shape restrictions on the option price function. The presence of arbitrage in the option price inverse problem may make the model break down [5, 31, 82]. So, when we use the

Black-Scholes formula and the interpolated volatilities to price options, we need to make sure that the final option price function satisfies Items 1 and 2.

Convexity and monotonicity of solutions to Dupire's partial differential equation (9) and (10) with respect to the strike were obtained in [22], as follows:

Theorem A.1.1. *[22, Theorem 4.3] The solution of Dupire's equation* (9) *with initial condition* (10) *is convex and non-increasing with respect to K, non-decreasing with respect to the local volatility and converges to* 0*, when $K \to \infty$, uniformly with T.*

A.2 Tikhonov Functional and its Minimum

Tikhonov regularization of the inverse problem for the volatility surface $\sigma(K, T)$ was proposed previously in Chapter 2. A regularized volatility surface was obtained by minimization of the Tikhonov functional (2.1), where f is a convex penalization. Other approaches to calibrate the local volatility by Tikhonov regularization were reviewed in Section 2.1.1. However, the question is: Is the calibrated volatility resulting of the minimization of the Tikhonov functional, an arbitrage-free volatility surface?

We show below that, in a market model where the dividend rates q is smaller than the interesting rates r, any algorithm that gives calibrate volatilities belongs to $\mathcal{D}(F)$ produces arbitrage-free option prices.

Lemma A.2.1. *Let $a \in \mathcal{D}(F)$. If $u(a)$ is a solution of* (9) *with initial condition* (10) *and $q \leq r$, then $u(a)$ satisfies the no-arbitrage conditions of Items 1 and 2.*

Proof. Note that, from Theorem A.1.1, we have that $U_K \leq 0$ and $U_{KK} \geq 0$. Using the change of variables (7), we conclude that $u_y \leq 0$. Moreover, from [58, Theorem 9.3.1] $u_{yy} - u_y > 0$. Positivity of $a \in \mathcal{D}(F)$ and definition of the equation (9) implies that $u_\tau > 0$. □

However, only convexity of the call option price U with respect to the strike K, can, in practice, imply in violation of the no-arbitrage restrictions. Indeed, as equation (A.2) shows, U_{KK} is equivalent to the risk-neutral density. Hence, we cannot choose a freely interpolation methods. Because of this, the model in [31] sometimes has negative probabilities on some nodes and [5] found negative local volatility problems. In order to overcome this problem, [82] proposes an interpolation method that consists in minimizing the distance between the density and the prior density in L^2. An often good choice for a log-normal, prior, density function is

$$\rho_0(K|S_t,\tau,r) = (1/K\sqrt{2\pi a_0\tau})\exp\left[-(\log(K)-\log(S_t)-r\tau+a_0\tau/2)^2/2a_0\tau\right] \tag{A.3}$$

with at-the-money option implies volatility $a_0 \in \mathcal{D}(F)$.

Using the approach derived in Section 2.2, we propose to combine the useful prior information to make the approximation more accurate.

According to Example 2.2.3, if we incorporate the assumption that the conditional probability density of observed data U_{KK} given $e^{-r\tau}\rho_0(K|S_t,\tau,r)$ is normally distributed, then we have the Tikhonov regularization given by

$$\mathcal{F}_{\alpha,\lambda}(a) := \frac{1}{2}\left\|F(a)-u^\delta\right\|^2_{L^2(\Omega)} + \alpha\{R_\lambda(a) + KL(a,a_0)\}\,, \tag{A.4}$$

where,

$$R_\lambda(a) := \lambda\left\|U_{KK}(a) - e^{-r\tau}\rho_0(K|S_t,\tau,r)\right\|^2_{L^2(\Omega)}\,.$$

Here, α plays the role of the regularization parameter and λ acts as a weight. The incorporation of the functional R_λ in the Tikhonov functional (A.4) can be interpreted as a data preprocessing step that adds some more regularity to our approach. In [82], the functional $R_\lambda(\cdot)$ is used to construct a B-spline interpolation method that preserves the shape restriction of the option price function.

Continuity, compactness and weakly closedness of $F : \mathcal{D}(F) \longrightarrow W_2^{1,2}(\Omega)$ shows that $R_\lambda(\cdot)$ is weakly lower semi-continuous, convex. Therefore, the convergence analysis in Chapter 2 is applicable to the Tikhonov functional (A.4).

Remark A.2.1. *Note that if we incorporate in Example 2.2.3 an exponential family associated to a Poisson distribution of observed $\rho(K|S_t,r,\tau)$ given ρ_0, we obtain a Tikhonov regularization with $R_\lambda(\rho,\rho_0) = KL(\rho,\rho_0)$.*

A particular Tikhonov functional with convex regularization is

$$\mathcal{F}_\alpha(a) := \frac{1}{2}\left\|F(a)-u^\delta\right\|^2_{L^2(\Omega)} + \alpha\|a-a_0\|^2_{\mathcal{H}(\Omega)}\,, \tag{A.5}$$

where α is the regularization parameter and $a_0 \in \mathcal{D}(F)$ is some *a priori* information about the true solution of the problem (15).

The minimizers of (A.5) generate arbitrage-free price surfaces.

In the next section, we look for convergence rates of this particular Tikhonov functional. Using properties of the problem, we are able to show better convergence rates than those proposed in [22, 35].

A.3 Convergence Rates and No-Arbitrage

We shall now focus on the convergence rate analysis of the minimizer for the Tikhonov functional (A.5).

In order to do this, we need some regularity assumptions on the a_0-minimal-norm solution $a^\dagger$ of (15). This assumption is called the source condition in regularization theory. A natural source condition to this problem is

$$a^\dagger - a_0 \in \mathcal{R}(F'(a^\dagger)^*)\,. \tag{A.6}$$

Remark A.3.1. *We remark that the adjoint of $F'(\cdot)$ in (A.6) is considered in the $L^2(\Omega)$-inner product. Indeed, we are looking for the continuous extension of $F'(\cdot)$ to $L^2(\Omega)$. A preliminary review of the density of $\mathcal{H}(\Omega)$ and $W_2^{1,2}(\Omega)$ in $L^2(\Omega)$ and the Banach extension theorem provides support for the claim. Hence, it follows from Lemma 2.1.1 that there exists $w^\dagger \in L^2(\Omega)$ and $r^\dagger = r(w^\dagger) \in L^2(\Omega)$ such that the approximate source condition*

$$\frac{1}{2}\partial\left\|a^\dagger - a_0\right\|^2 = a^\dagger - a_0 = F'(a^\dagger)^* w^\dagger + r^\dagger \tag{A.7}$$

is attainable. Moreover, for a fixed given $\delta > 0$, it follows that $\left\|r^\dagger\right\|_{L^2(\Omega)} \leq \delta$.

From the existence of an approximate source condition (A.7), we are able to obtain convergence rates of the regularized solutions that were obtained by minimizing the Tikhonov functional (A.5).

Theorem A.3.1. *Let $a^\dagger \in \mathcal{D}(F)$ denote an a_0-minimum-norm-solution of $F(a) = u(a)$ and a_α^δ the minimizer of (A.5), with u^δ satisfying equation (13). Suppose that*

$$C\left\|w^\dagger\right\| < 1/2\,, \tag{A.8}$$

where C is the constant in (1.13) and $w^\dagger$ given in Remark A.3.1. From the choice of $\alpha \sim \delta$, *we obtain*

$$\left\|a_{\alpha^\delta} - a^\dagger\right\|_{\mathcal{H}(\Omega)} = \mathcal{O}(\delta) \quad \textit{and} \quad \left\|F(a_\alpha^\delta) - u^\delta\right\| = \mathcal{O}(\delta). \tag{A.9}$$

Proof. Let a_α^δ be a minimizer of $\mathcal{F}_\alpha$. The definition of $\mathcal{F}_\alpha$ and (13) implies that

$$\left\|F(a_\alpha^\delta) - u^\delta\right\|_{L^2(\Omega)}^2 + \alpha\left\|a_\alpha^\delta - a_0\right\|_{\mathcal{H}(\Omega)}^2 \le \mathcal{F}_\alpha(a_\alpha^\delta) \le \mathcal{F}_\alpha(a^\dagger) \le \delta^2 + \alpha\left\|a^\dagger - a_0\right\|_{\mathcal{H}(\Omega)}^2 .$$

Since, $\left\|a^\dagger - a_\alpha^\delta\right\|_{\mathcal{H}(\Omega)}^2 = \left\|a_\alpha^\delta - a_0\right\|_{\mathcal{H}(\Omega)}^2 - \left\|a^\dagger - a_0\right\|_{\mathcal{H}(\Omega)}^2 - 2\langle a^\dagger - a_0, a_\alpha^\delta - a^\dagger\rangle$, we have from the above estimate and (A.7) that

$$\begin{aligned}
\left\|F(a_\alpha^\delta) - u^\delta\right\|_{L^2(\Omega)}^2 + \alpha\left\|a_\alpha^\delta - a^\dagger\right\|_{\mathcal{H}(\Omega)}^2 &\le \delta^2 - 2\alpha\langle a^\dagger - a_0, a_\alpha^\delta - a^\dagger\rangle \\
&= \delta^2 - 2\alpha\langle F'(a^\dagger)^* w^\dagger + r^\dagger, a_\alpha^\delta - a^\dagger\rangle \\
&= \delta^2 + 2\alpha\langle w^\dagger, -F'(a^\dagger)(a_\alpha^\delta - a^\dagger)\rangle - 2\alpha\langle r^\dagger, a_\alpha^\delta - a^\dagger\rangle \\
&\le \delta^2 + 2\alpha\left\|w^\dagger\right\|\left\|F(a_\alpha^\delta) - F(a^\dagger) - F'(a^\dagger)(a_\alpha^\delta - a^\dagger)\right\| \\
&\qquad + 2\alpha\left\|w^\dagger\right\|\left\|F(a_\alpha^\delta) - F(a^\dagger)\right\| + 2\alpha\left\|r^\dagger\right\|\left\|a_\alpha^\delta - a^\dagger\right\| .
\end{aligned}$$

Using (1.13), it follows that

$$\left\|F(a_\alpha^\delta) - F(a^\dagger) - F'(a^\dagger)(a_\alpha^\delta - a^\dagger)\right\| \le C\left\|a_\alpha^\delta - a^\dagger\right\|^2 .$$

From the last two inequalities we get

$$\begin{aligned}
&\|F(a_\alpha^\delta) - u^\delta\|_{L^2(\Omega)}^2 + \alpha\left\|a_\alpha^\delta - a^\dagger\right\|_{\mathcal{H}(\Omega)}^2 \\
&\qquad \le \delta^2 + 2C\alpha\left\|w^\dagger\right\|\left\|a_\alpha^\delta - a^\dagger\right\|^2 + 2\alpha\delta\left\|w^\dagger\right\| + 2\alpha\left\|w^\dagger\right\|\left\|F(a_\alpha^\delta) - u^\delta\right\| + 2\alpha\left\|r^\dagger\right\|\left\|a_\alpha^\delta - a^\dagger\right\| .
\end{aligned}$$

Using (A.8) and Lemma 2.1.1, we conclude that

$$\begin{aligned}
\left(\left\|F(a_\alpha^\delta) - u^\delta\right\|_{L^2(\Omega)} - \alpha\left\|w^\dagger\right\|\right)^2 + \alpha(1 - 2C\left\|w^\dagger\right\|)\left(\left\|a_\alpha^\delta - a^\dagger\right\|_{\mathcal{H}(\Omega)} - \left\|r^\dagger\right\|/(1 - 2C\left\|w^\dagger\right\|)\right)^2 \\
\le \delta^2 + 2\alpha\delta\left\|w^\dagger\right\| + \alpha^2\left\|w^\dagger\right\|^2 + \alpha\left\|r^\dagger\right\|^2/(1 - 2C\left\|w^\dagger\right\|) \\
\le (\delta + \alpha\left\|w^\dagger\right\|)^2 + \alpha\delta^2/(1 - 2C\left\|w^\dagger\right\|) .
\end{aligned}$$

Now, with the choice of $\alpha \sim \delta$, we have the assertion. □

Note that, in the above theorem, we improve the convergence rates of the regularized solution without using the general framework developed in Chapter 2.

Appendix B

Background Material and Definitions

In this appendix we present some definitions and concepts used in this work. In the sequel, U and V denote Banach spaces and U^* and V^* their respective dual spaces.

B.1 Convex Analysis

We start by recalling some definitions from convex analysis [19, 21, 38] used throughout this work.

Definition 6. *A subset $\tilde{U}$ of U is convex if, for any $a, b \in \tilde{U}$, we have*

$$\alpha a + (1-\alpha)b \in \tilde{U} \qquad , \forall \alpha \in [0,1]\,.$$

Definition 7. *A functional $f : U \longrightarrow \mathbb{R} \cup \{+\infty\}$ is said to be a convex function if the domain of f*

$$\mathcal{D}(f) := \{a \in U \,:\, f(a) < +\infty\}$$

is a convex subset of U and

$$f(\alpha a + (1-\alpha)b) \leq \alpha f(a) + (1-\alpha) f(b) \quad \forall a, b \in \mathcal{D}(f) \quad \text{and} \quad \forall \alpha \in [0,1]\,.$$

A convex function f is called proper *if $\mathcal{D}(f) \neq \emptyset$.*

Definition 8. *Let $f : U \longrightarrow \mathbb{R}$ be a real-valued convex function defined in a convex set $\tilde{U}$ of a Banach space U. An element $\xi \in U^*$ is called a called a subgradient at a point $a^\dagger$ in*

$\tilde{U}$ if for any a in $\tilde{U}$ one has

$$f(a) - f(a^\dagger) \geq \langle \xi, a - a^\dagger \rangle \, . \tag{B.1}$$

Definition 9. *The set of all subgradients at $a^\dagger$ is called the subdifferential at $a^\dagger$ and is denoted $\partial f(a^\dagger)$.*

An important tool in the study of Tikhonov regularization in Chapter 2 is the notion of Bregman distance.

Definition 10. *Let f be a convex function. For given $a \in \mathcal{D}(f)$, let $\partial f(a) \subset U^*$ denote the subdifferential of the functional f at a. We denote by*

$$\mathcal{D}(\partial f) = \{\tilde{a} : \partial f(\tilde{a}) \neq \emptyset\}$$

the domain of the subdifferential [21]. The Bregman distance with respect to $\zeta \in \partial f(a_1)$ is defined on $\mathcal{D}(f) \times \mathcal{D}(\partial f)$ by

$$D_\zeta(a_2, a_1) = f(a_2) - f(a_1) - \langle \zeta, a_2 - a_1 \rangle \, .$$

It is important to notice that the Bregman distance is not a distance in the actual meaning of the term, since it is not necessarily symmetric. It is however always positive and if f is proper and strictly convex it vanishes if and only if $a_1 = a_2$.

The concept of Fenchel conjugation can be motivated as a generalization of the Legendre transform in Hamiltonian mechanics. It plays an important role in the connection with Convex Risk Measures in Chapter 4.

Definition 11. *The Fenchel conjugation of a convex function $f : U \longrightarrow \mathbb{R} \cup \{+\infty\}$ is the function $f^* : U^* \longrightarrow \mathbb{R} \cup \{+\infty\}$ defined by*

$$f^*(\phi) = \sup_{a \in U} \{\langle \phi, a \rangle - f(a)\} \, ,$$

where $\langle \cdot, \cdot \rangle$ denotes the duality application between U and U^*.

Note that, f^* is convex. Moreover, if $\mathcal{D}(f) \neq \emptyset$, then f^* never take the value $-\infty$.

B.2 Sufficient Statistic and Exponential Families

We briefly recall the concept of exponential families and sufficient statistic. Such concepts play a fundamental rôle in a number of areas. In particular, from our perspective the connection between inverse problems and information geometry. See [3, 4, 75] and references therein.

Consider a measurable space $(\Omega, \mathcal{B})$ with $\mathcal{B}$ a σ-algebra of subsets of Ω. Let $\mathbf{t}$ a measurable function from Ω into $\mathbb{R}$. [1]

Let $p_0 : \mathbb{R} \longrightarrow \mathbb{R}_+$ be any function such that if $(\Omega, \mathcal{B})$ is endowed with a measure $dP_0(\omega) = p_0(\mathbf{t}(\omega))d\mathbf{t}(\omega)$, then $\int_{\omega\in\Omega} dP_0(\omega) < \infty$. Note that the measure P_0 is absolutely continuous with respect to the Lebesgue measure $d\mathbf{t}(\omega)$.

Thus, $\mathbf{t}(\omega)$ is a random variable from $(\Omega, \mathcal{B}, P_0)$ to $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, where $\mathcal{B}(\mathbb{R})$ denotes de σ-algebra of Borel sets in $\mathbb{R}$.

Let Θ be defined as the set of all parameters $\theta \in \mathbb{R}$ for which

$$\int_{\omega\in\Omega} \exp\left[\langle\theta, \mathbf{t}(\omega)\rangle\right] dP_0(\omega) < \infty\,.$$

Define the function $\psi : \Theta \longrightarrow \mathbb{R}$ such that

$$\psi(\theta) = \log\left(\int_{\omega\in\Omega} \exp\left[\langle\theta, \mathbf{t}(\omega)\rangle\right] dP_0(\omega)\right)\,. \tag{B.2}$$

A family of probability distributions $\mathfrak{F}_\psi$ parameterized by $\theta \in \Theta$ such that the probability density function with respect to the measure $\mathbf{t}(\omega)$ can be expressed in the form

$$q(\omega;\theta) = \exp\left(\langle\theta, \mathbf{t}(\omega)\rangle - \psi(\theta)\right) p_0(\mathbf{t}(\omega))$$

is called an *exponential family with natural statistic* $\mathbf{t}(\omega)$*, natural parameter* θ *and natural parameter space* Θ. If, in addition, the parameter space Θ is an open set, then $\mathfrak{F}_\psi$ is called a *regular exponential family.*

Let s denote the natural statistic $\mathbf{t}(\omega)$. If the probability density (with respect to the appropriate measure ds) given by

$$g(s;\theta) = \exp\left[\langle\theta, s\rangle - \psi(\theta)\right] p_0(s)$$

[1] To fix ideas, the function $\mathbf{t}$ can be defined from Ω to $\mathbb{R}^N$.

is such that

$$\frac{q(\omega;\theta)}{g(s;\theta)}$$

does not depend on θ, then s is a *sufficient statistic* for the family.

Remark B.2.1. *For our analysis, we redefine the concept of regular exponential families in terms of the probability density s. But, the original probability space could be quite general.*

Definition 12. *A parametric family $\mathfrak{F}_\psi$ of distributions $\{p_{(\psi,\theta)} : \theta \in \Theta = \mathcal{D}(\psi)\}$ is called a regular exponential family if each probability density is of the form*

$$p_{(\psi,\theta)}(s) = \exp\left[\langle s, \theta\rangle - \psi(\theta)\right] p_0(s)\,, \tag{B.3}$$

where s is a sufficient statistic for the family.
The function $\psi(\theta)$ is known as the log partition function corresponding to the family.

B.3 Entropy

The concept of *entropy* is a measure of how organized a physical system is. It goes back to the principles of Thermodynamics and in particular to the *Second Law*, which states that the entropy of a system cannot decrease other than by increasing the entropy of another system. Statistical mechanics views entropy as the amount of uncertainty which remains about a system, after its observable macroscopic properties have been taken into account. More specifically, entropy is a logarithmic measure of the density of states:

$$g(p_i) = -\sum_{i \in I} p_i \log(p_i)\,, \tag{B.4}$$

where the summation is over all the microstates the system can be in, and the p_i are the probabilities for the system to be in the i-th microstate.

In information theory, entropy is the measure of the amount of information that is missing before reception. Sometimes referred to as Shannon entropy.

The definition in Equation (B.4) makes the tacit assumption the underlying measure is the counting measure on the discrete space I. In other words, it is a *relative* entropy with respect to the uniform distribution. As it turns out, in more general contexts, it is necessary to consider the relative entropy of one measure with respect to another. This brings in the concept of the Kullback-Leibler distance between two probability density functions. More specifically, we have for two probability density functions p and q the following natural definition

$$\mathrm{KL}(p, q) = \int p(x) \log(q(x)/p(x))dx \ .$$

Of more generally if P and Q are the probability measures, with Q absolutely continuous with respect to P,

$$\mathrm{KL}(dP||dQ) := \int \log(dQ/dP)dP \ .$$

This can be rewritten as

$$\mathrm{KL}(dP||dQ) = \mathbb{E}^P \log(dQ/dP) \ .$$

Jensen's inequality yields that $\mathrm{KL}(dP||dQ) \geq 0$.

The concept of exponential family arises naturally in order to answer the question: What is the maximum entropy distribution consistent with given constraints on expected values?

More precisely, if we now consider a collection of random variables X_i. Then, the probability distribution P whose entropy with respect to Q is maximal, subject to the restriction that the expected value of X_i are equal to s_i, is a member of the exponential family with P as reference measure and $(X_1, ..., X_n)$ as sufficient statistic.

Bibliography

[1] R. A. Adams, *Sobolev spaces*, Academic Press [A subsidiary of Harcourt Brace Jovanovich, Publishers], New York-London, 1975, Pure and Applied Mathematics, Vol. 65.

[2] Y. Altun and A. Smola, *Unifying divergence minimization and statistical inference via convex duality*, Learning theory, Lecture Notes in Comput. Sci., vol. 4005, Springer, Berlin, 2006, pp. 139–153.

[3] Shun-ichi Amari and Hiroshi Nagaoka, *Methods of information geometry. Transl. from the Japanese by Daishi Harada.*, Translations of Mathematical Monographs. 191. Providence, RI: American Mathematical Society (AMS). Oxford: Oxford University Press., 2000 (English).

[4] E. B. Andersen, *Sufficiency and exponential families for discrete sample spaces*, **65** (1970), 1248–1255.

[5] L. Andersen and R. Brotherton-Ratcliffe, *The equality option volatility smile: An implicit finite-difference approach*, Journal of Computational Finance **1** (1998), 5–38.

[6] P. Artzner, F. Delbaen, J-M. Eber, and D. Heath, *Coherent measures of risk*, Math. Finance **9** (1999), no. 3, 203–228.

[7] M. Avellaneda, C. Fridman, R. Holmes, and D. Samperi, *Calibrating volatility surfaces via relative entropy minimization*, Applied Mathematical Finance **4** (1997), 37–64.

[8] A. Banerjee, S. Merugu, I.S. Dhillon, and J. Ghosh, *Clustering with bregman divergences*, Journal of Machine Learning Research **6** (2005), 1705–1749.

[9] D. Bates, *Jumps and stochastic volatility: Exchange rate processes implicit in deutsche mark options*, Rev. of Finance Studies **9** (1996), 69–107.

[10] J. Baumeister, *Stable solution of inverse problems*, Advanced Lectures in Mathematics, Friedr. Vieweg & Sohn, Braunschweig, 1987.

[11] J. Baumeister, A. De Cezaro, and A. Leitao, *Modified iterated Tikhonov methods for solving systems of nonlinear ill-posed equations*, (2009), submited.

[12] J. Baumeister, B. Kaltenbacher, and A. Leitao, *On Levemberg-Marquardt Kaczmarz methods for regularizing systems of nonlinear ill-posed equations*, Inverse Problems and Imaging (2010), no. 4, to appear.

[13] J. Baumeister and A. Leitão, *Topics in inverse problems*, Publicações Matemáticas do IMPA. [IMPA Mathematical Publications], Instituto Nacional de Matemática Pura e Aplicada (IMPA), Rio de Janeiro, 2005, 25o Colóquio Brasileiro de Matemática. [25th Brazilian Mathematics Colloquium].

[14] A. Binder, M. Hanke, and O. Scherzer, *On the Landweber iteration for nonlinear ill-posed problems*, J. Inverse Ill-Posed Probl. **4** (1996), no. 5, 381–389, 2nd International Symposium "Inverse Problems in Engineering Sciences" (Osaka, 1994).

[15] F. Black and M. Scholes, *The pricing of options and corporate liabilities*, J. Political Economy **81** (1973), 637–659.

[16] I. Bouchouev and V. Isakov, *Uniqueness, stability and numerical methods for the inverse problem that arises in financial markets*, Inverse Problems **15** (1999), no. 3, R95–R116.

[17] M. Burger and B. Kaltenbacher, *Regularizing Newton-Kaczmarz methods for nonlinear ill-posed problems*, SIAM J. Numer. Anal. **44** (2006), no. 1, 153–182 (electronic).

[18] M. Burger and S. Osher, *Convergence rates of convex variational regularization*, Inverse Problems **20** (2004), no. 5, 1411–1421.

[19] D Butnariu and A. N. Iusem, *Totally convex functions for fixed points computation and infinite dimensional optimization*, Applied Optimization, vol. 40, Kluwer Academic Publishers, Dordrecht, 2000.

[20] R. Carmona and S. Nadtochiy, *Local volatility dynamic models*, Finance Stoch. **13** (2009), 1–48.

[21] F.H. Clarke, *Optimization and Nonsmooth Analysis*, Wiley-Interscience, Vancouver, 1983, 1th edition.

[22] S. Crépey, *Calibration of the local volatility in a generalized Black-Scholes model using Tikhonov regularization*, SIAM J. Math. Anal. **34** (2003), no. 5, 1183–1206 (electronic).

[23] ———, *Calibration of the local volatility in a trinomial tree using Tikhonov regularization*, Inverse Problems **19** (2003), no. 1, 91–127.

[24] A. De Cezaro, M. Hiltermiler, A. Leitao, and O. Scherzer, *On steepest-descent-Kaczmarz methods for regularizing systems of nonlinear ill-posed equtions*, Applied Mathematics and Computation **202** (2008), 596–607.

[25] A. De Cezaro, A. Leitão, and X.-C. Tai, *On a level-set type methods for recovering picewise constant solution of ill-posed problems*, Lecture Notes in Computer Science (2009), no. 5567, 50–62.

[26] ———, *On multiple level-set regularization methods for inverse problems*, Inverse Problems **25** (2009), no. 035004, 22p.

[27] ———, *Level-set of L^2type for recovering shape and contrast in inverse problems*, submited (2010), 1–20.

[28] A. De Cezaro, O. Scherzer, and J.P. Zubelli, *Convex regularization applied to the inverse problem of option pricing: convergence analysis and rates*, submitted (2010), 1 – 30.

[29] ———, *A convex-regularization framework for local-volatility calibration in derivative markets: The connection with convex risk measures and exponential families*, 6th World Congress of the Bachelier Finance Society **accepted** (2010), 1 – 19.

[30] A. De Cezaro and J.P. Zubelli, *Iterative calibration of volatility surface*, **working paper** (2010).

[31] E. Derman, I. Kani, and N. Chirss, *Implied trimonial trees of the volatility smile*, Journal of Derivatives **3** (1996), 7–22.

[32] J. Diestel, *Geometry of Banach spaces—selected topics*, Lecture Notes in Mathematics, Vol. 485, Springer-Verlag, Berlin, 1975.

[33] B. Dupire, *Pricing with a smile*, Risk **7** (1994), 18– 20.

[34] B. Dupire, *Pricing and hedging with smiles*, Mathematics of derivative securities (Cambridge, 1995), Publ. Newton Inst., vol. 15, Cambridge Univ. Press, Cambridge, 1997, pp. 103–111.

[35] H. Egger and H. W. Engl, *Tikhonov regularization applied to the inverse problem of option pricing: convergence analysis and rates*, Inverse Problems **21** (2005), no. 3, 1027–1045.

[36] H. Egger, T. Hein, and B. Hofmann, *On decoupling of volatility smile and term structure in inverse option pricing*, Inverse Problems **22** (2006), no. 4, 1247–1259.

[37] P. P. B. Eggermont, *Maximum entropy regularization for Fredholm integral equations of the first kind*, **24** (1993), no. 6, 1557–1576.

[38] I. Ekeland and R. Temam, *Convex analysis and variational problems*, North-Holland, Amsterdam, 1976.

[39] H. W. Engl, M. Hanke, and A. Neubauer, *Regularization of inverse problems*, Mathematics and its Applications, vol. 375, Kluwer Academic Publishers Group, Dordrecht, 1996.

[40] H. W. Engl and G. Landl, *Convergence rates for maximum entropy regularization*, SIAM J. Numer. Anal. **30** (1993), no. 5, 1509–1536.

[41] Heinz W. Engl and William Rundell (eds.), *Inverse problems in diffusion processes*, Philadelphia, PA, Society for Industrial and Applied Mathematics (SIAM), 1995.

[42] T. Fischer, *Risk capital allocation by coherent risk measures based on one-sided moments*, Insurance Math. Econom. **32** (2003), no. 1, 135–146.

[43] H. Föllmer and A. Schied, *Convex measures of risk and trading constraints*, Finance Stoch. **6** (2002), no. 4, 429–447.

[44] ______, *Robust preferences and convex measures of risk*, Advances in finance and stochastics, Springer, Berlin, 2002, pp. 39–56.

[45] E. Fournié, J.-M. Lasry, J. Lebuchoux, P.-L. Lions, and N. Touzi, *Applications of Malliavin calculus to Monte Carlo methods in finance*, Finance Stoch. **3** (1999), no. 4, 391–412.

[46] A. Friedman, *Variational principles and free-boundary problems*, Pure and Applied Mathematics, John Wiley & Sons Inc., New York, 1982, A Wiley-Interscience Publication.

[47] M. Grasmair, M. Haltmeier, and O. Scherzer, *Sparse regularization with l^q penalty term*, Inverse Problems **24** (2008), no. 5, 055020, 13.

[48] J. Hadamard, *Lectures on cauchy's problem in linear partial differential equations*, Dover Publications, 1953.

[49] M. Haltmeier, R. Kowar, A. Leitão, and O. Scherzer, *Kaczmarz methods for regularizing nonlinear ill-posed equations. II. Applications*, Inverse Probl. Imaging **1** (2007), no. 3, 507–523.

[50] M. Haltmeier, A. Leitão, and O. Scherzer, *Kaczmarz methods for regularizing nonlinear ill-posed equations. I. Convergence analysis*, Inverse Probl. Imaging **1** (2007), no. 2, 289–298.

[51] M. Hanke, A. Neubauer, and O. Scherzer, *A convergence analysis of Landweber iteration for nonlinear ill-posed problems*, Numer. Math. **72** (1995), 21–37.

[52] T. Hein and B. Hofmann, *On the nature of ill-posedness of an inverse problem arising in option pricing*, Inverse Problems **19** (2003), no. 6, 1319–1338.

[53] ———, *Approximate source conditions for nonlinear ill-posed problems—chances and limitations*, Inverse Problems **25** (2009).

[54] B. Hofmann, B. Kaltenbacher, C. Pöschl, and O. Scherzer, *A convergence rates result for Tikhonov regularization in Banach spaces with non-smooth operators*, Inverse Problems **23** (2007), no. 3, 987–1010.

[55] B. Hofmann and R. Krämer, *On maximum entropy regularization for a specific inverse problem of option pricing*, J. Inverse Ill-Posed Probl. **13** (2005), no. 1, 41–63.

[56] J. Hull and A. White, *The pricing of options with sthochastic volatility*, Journal of Finance **42** (1987), 281–300.

[57] V. Isakov, *Inverse source problems*, Mathematical Surveys and Monographs, vol. 34, American Mathematical Society, Providence, RI, 1990.

[58] ———, *Inverse problems for partial differential equations*, second ed., Applied Mathematical Sciences, vol. 127, Springer, New York, 2006.

[59] N. Jackson, E. Süli, and S. Howison, *Computation of deterministic volatility surfaces*, Journal Comp. Finance **2** (1998), 5–32.

[60] S. Kaczmarz, *Approximate solution of systems of linear equations*, Internat. J. Control **57** (1993), no. 6, 1269–1271, Translated from the German.

[61] B. Kaltenbacher, A. Neubauer, and O. Scherzer, *Iterative regularization methods for nonlinear ill-posed problems*, Radon Series on Computational and Applied Mathematics, vol. 6, Walter de Gruyter GmbH & Co. KG, Berlin, 2008.

[62] N. Karalé, *Smile interpolation and calibration of the local volatility model*, http://nkahale.free.fr/papers/Interpolation.pdf (2005), 1–18.

[63] I. Karatzas and S. E. Shreve, *Methods of mathematical finance*, Applications of Mathematics (New York), vol. 39, Springer-Verlag, New York, 1998.

[64] A. Kirsch, *An introduction to the mathematical theory of inverse problems*, Applied Mathematical Sciences, vol. 120, Springer-Verlag, New York, 1996.

[65] R. Korn and E. Korn, *Option pricing and portfolio optimization*, Graduate Studies in Mathematics, vol. 31, American Mathematical Society, Providence, RI, 2001, Modern methods of financial mathematics, Translated from the 1999 German original by the authors.

[66] R. Kowar and O. Scherzer, *Convergence analysis of a Landweber-Kaczmarz method for solving nonlinear ill-posed problems*, Ill-posed and inverse problems, VSP, Zeist, 2002, pp. 253–270.

[67] S. Kullback and R. A. Leibler, *On information and sufficiency*, Ann. Math. Statistics **22** (1951), 79–86.

[68] R. Lagnardo and S. Osher, *A technique for calibtrating derivative security pricing models: numerical solution of an inverse problem*, Journal of Computational Finance **1** (1997), no. 1, 13–25.

[69] L. Landweber, *An iteration formula for Fredholm integral equations of the first kind*, Amer. J. Math. **73** (1951), 615–624.

[70] O. A. Landyzenskaya, V. A. Solonikov, and N. N. Urealceva, *Linear and quasilinear equations of parabolic type*, Translations of Mathematical Monographs vol 23, AMS, Providence, RI, 1968.

[71] A. Meucci, *Risk and asset allocation*, Springer Finance, Springer-Verlag, Berlin, 2005.

[72] F. Natterer and F. Wübbeling, *Mathematical Methods in Image Reconstruction*, SIAM, Philadelphia, 2001.

[73] D. Nualart, *The Malliavin calculus and related topics*, Probability and its Applications (New York), Springer-Verlag, New York, 1995.

[74] E. Resmerita, *Regularization of ill-posed problems in Banach spaces: convergence rates*, Inverse Problems **21** (2005), no. 4, 1303–1314.

[75] E. Resmerita and R. S. Anderssen, *Joint additive Kullback–Leibler residual minimization and regularization for linear inverse problems*, **30** (2007), no. 13, 1527–1544.

[76] E. Resmerita and O. Scherzer, *Error estimates for non-quadratic regularization and the relation to enhancement*, Inverse Problems **22** (2006), no. 3, 801–814.

[77] R. T. Rockafellar, *Conjugate duality and optimization*, Society for Industrial and Applied Mathematics, Philadelphia, Pa., 1974, Lectures given at the Johns Hopkins University, Baltimore, Md., June, 1973, Conference Board of the Mathematical Sciences Regional Conference Series in Applied Mathematics, No. 16.

[78] O. Scherzer, *A convergence analysis of a method of steepest descent and a two-step algorithm for nonlinear ill-posed problems*, Numer. Funct. Anal. Optim. **17** (1996), no. 1-2, 197–214.

[79] O. Scherzer, M. Grasmair, H. Grossauer, M. Haltmeier, and F. Lenzen, *Variational methods in imaging*, Applied Mathematical Sciences, vol. 167, Springer, New York, 2008.

[80] M. Schweizer and J. Wissel, *Term structure implied volatilies: Absence arbitrage and existence results*, Mathematical Finance **18** (2008), no. 1, 77–114.

[81] A. N. Tikhonov and V. Y. Arsenin, *Solutions of ill-posed problems*, John Wiley & Sons, Washington, D.C., 1977, Translation editor: Fritz John.

[82] Y. Wang, H. Yin, and L. Qi, *No-arbitrage interpolation of the option price function and its reformulation*, J. Optim. Theory Appl. **120** (2004), no. 3, 627–649.

[83] P. Wilmott and J. Howison, S.and Dewynne, *The mathematics of financial derivatives*, Cambridge University Press, Cambridge, 1995, A student introduction.

[84] K. Yosida, *Functional analysis*, Classics in Mathematics, Springer-Verlag, Berlin, 1995, Reprint of the sixth (1980) edition.

[85] C. Zălinescu, *Convex analysis in general vector spaces*, World Scientific Publishing Co. Inc., River Edge, NJ, 2002.

[86] E. Zeidler, *Nonlinear functional analysis and its applications. I*, Springer-Verlag, New York, 1986, Fixed-point theorems, Translated from the German by Peter R. Wadsack.

[87] Jorge P. Zubelli, *An introduction to inverse problems. Examples, methods and questions*, 22º Colóquio Brasileiro de Matemática. [22nd Brazilian Mathematics Colloquium], Instituto de Matemática Pura e Aplicada (IMPA), Rio de Janeiro, 1999, Appendix A by the author and Luis Orlando Castellano Pérez.