MASTERARBEIT | MASTER'S THESIS

Titel | Title

Hedging Under Uncertainty

verfasst von | submitted by

Afonso José Andrade Luís

angestrebter akademischer Grad | in partial fulfilment of the requirements for the degree of

Master of Science (MSc)

Wien | Vienna, 2026

Studienkennzahl lt. Studienblatt | Degree programme code as it appears on the student record sheet:	UA 066 821
Studienrichtung lt. Studienblatt | Degree programme as it appears on the student record sheet:	Masterstudium Mathematik
Betreut von | Supervisor:	Assoz. Prof. Dr. Julio Daniel Backhoff-Veraguas

Acknowledgements

First and foremost, I would like to express my deepest gratitude to my family, without whom the undertaking of this Master's degree, culminating in this thesis, would not have been possible. I am forever grateful to them not only for giving me the opportunity to study abroad, but also for their constant support, undying belief in me and words of encouragement: something I find difficult to put into words, but which I will always carry dearly and warmly in my heart. To my father, David Fernandes Luís; my mother, Maria Idalina de Branco Andrade Luís; and my brother, João Pedro Andrade Luís, I dedicate this thesis equally: Obrigado por tudo!

Secondly, I would like to thank the friends I have made here in Vienna. Without them as well, this thesis would not have been possible. I am grateful for the unforgettable experiences they have shared with me, which have certainly enriched me as a person and made my stay abroad much easier and more enjoyable.

Lastly, and no less importantly, I would like to express my appreciation to my supervisor, Professor Julio. I am forever thankful for his boundless patience, his ever-present willingness to help, and his mathematical rigor and guidance throughout the entirety of this thesis.

Abstract

The thesis is divided into two main parts. The first develops the theoretical setting of stochastic differential equations, including the notions of strong and weak solutions, path-wise uniqueness, and the characterization of strong solutions. Particular attention is devoted to existence and uniqueness theorems under global and local Lipschitz conditions, as well as to the $C^{\infty}-$ version of the diffeomorphism theorem and the non-crossing lemma. Building on this probabilistic foundation, the second part focuses in mathematical finance, with particular emphasis on the robustness of the Black–Scholes formula under volatility misspecification. After revisiting the classical pricing and hedging of an European call option under the Black–Scholes model, we investigate the impact of volatility misspecification on option prices and the associated hedging strategies. We then study an optimization problem in which a family of models is penalized, through an uncertainty-aversion parameter ψ, relative to a reference local volatility model. In particular, in the limit of small uncertainty aversion, explicit asymptotic formulas for the prices and hedging strategies of vanilla options are derived. Finally, we consider a novel multiplicative setting with logarithmic utility, where explicit expressions for the optimal hedging strategies and corresponding prices are also obtained.

Abstrakt

Die vorliegende Arbeit gliedert sich in zwei Hauptteile. Im ersten Teil wird der theoretische Rahmen stochastischer Differentialgleichungen entwickelt. Dabei werden die Begriffe starker und schwacher Lösungen, pfadweise Eindeutigkeit sowie die Charakterisierung starker Lösungen eingeführt und untersucht. Besonderes Augenmerk liegt auf Existenz- und Eindeutigkeitssätzen unter globalen und lokalen Lipschitz-Bedingungen sowie auf einer Version des Diffeomorphiesatzes und dem Nicht-Kreuzungs-Lemma. Aufbauend auf diesen probabilistischen Grundlagen widmet sich der zweite Teil der Finanzmathematik, insbesondere der Robustheit der Black–Scholes-Formel bei fehlspezifizierter Volatilität. Nach einer Darstellung der klassischen Bewertung und Absicherung einer europäischen Kaufoption im Black–Scholes-Modell untersuchen wir die Auswirkungen einer fehlerhaften Volatilitätsannahme auf Optionspreise und die zugehörigen Hedging-Strategien. Anschließend betrachten wir ein Optimierungsproblem, in dem eine Modellfamilie mittels eines Unsicherheitsaversion-Parameters relativ zu einem Referenzmodell mit lokaler Volatilität penalisiert wird. Insbesondere werden im Grenzfall geringer Unsicherheitsaversion explizite asymptotische Formeln für die Preise und Hedging-Strategien von Vanilla-Optionen hergeleitet. Abschließend untersuchen wir einen neuartigen multiplikativen Ansatz mit logarithmischer Nutzenfunktion, für den ebenfalls explizite Ausdrücke für die optimalen Hedging-Strategien und die entsprechenden Preise gewonnen werden.

Contents

1 Introduction

Mathematical finance relies fundamentally on stochastic modeling. Financial quantities such as asset prices, interest rates, and volatility are typically described through stochastic processes whose dynamics evolve under uncertainty. Among the mathematical tools used to model such phenomena, stochastic differential equations play a central role, providing a flexible framework capable of incorporating randomness directly into dynamical systems. The classical theory of option pricing and hedging, most notably the Black–Scholes model, is formulated in terms of diffusion processes driven by Brownian motion. Under suitable assumptions, these models allow for explicit pricing formulas and perfect replication strategies. However, many of the assumptions underlying these classical models are difficult to justify empirically. In particular, the precise specification of volatility is often uncertain, and even small misspecifications may lead to significant hedging errors. This motivates the study of hedging problems under model uncertainty, where one seeks strategies that remain effective across multiple possible probabilistic models.

In section (2), we study the theory of stochastic differential equations in the strong formulation. We introduce the canonical path space (2), predictable path functionals (3), and the notion of solutions relative to a prescribed stochastic basis (1). We focus on defining pathwise uniqueness (4) and exactness, emphasizing the deterministic character of strong solutions (5), once the driving Brownian motion is fixed. We then establish existence and uniqueness results under global and local Lipschitz assumptions ((1) and (2) respectively), using Picard iteration arguments and localization techniques. The discussion is further extended to a rigorous characterization of strong solutions, culminating in the $C^{\infty}-$version of the Diffeomorphism Theorem (7) for stochastic flows and the Non-Crossing Lemma (4). We additionally state a weaker C^k- version of the Diffeomorphism Theorem (8). Moreover, we delve into the realm of Weak solutions, where we provide a definition of of a weak solution to an SDE (11) and, in accordance with the path-wise uniqueness definition within the strong formulation, give a proper definition of uniqueness in law (12). Finally, we state without proof the famous Yamada and Watanabe Theorem (9), which links the exactness of an SDE with both weak and strong formulations.

In section (3), we turn at once to applications in mathematical finance. We first give account of some definitions that will form the foundations of the remaining thesis. We revisit the classical Black–Scholes setting and analyze the replication of an European call, in order to obtain a first solid intuition on the whole setting of self-financing strategies and correspondant hedging strategies in complete markets. We then proceed to define the problem clearly: an agent wrongfully assumes that the underlying asset is governed by $\bar{\sigma}(t, S_t)$, while it is governed by true volatility σ_t. It is worth here to mention that we assume that $\bar{\sigma}$ depends randomly only on the asset S_t (hence our choice to present it explicitly as often as possible), for otherwise it would lead to counter-intuitive effects, as one can see in [6, Section 4, Theorem 4.1]. Indeed, in the latter case, [6] shows that an increase (decrease) of volatility may lead to a decrease (increase) of the price of an European call - a scenario with little resemblance in a practical scenario. Under this estabilished setting, we proceed to prove the convexity of European contingent claims (6), which then proves to be useful in the proof of the main theorem of this section (10). Within the latter, we finally end up with mathematical evidence of the robustness of the Black-Scholes Model, as it shows very clearly that, assuming that we have an almost sure dominance of $(\sigma(t, S_t))_{t\geqslant 0}$ over $(\bar{\sigma}(t, S_t))_{t\geqslant 0}$ a.s), our portfolio at the terminal time (which is constructed according to the hedging strategy identified by $\bar{\sigma}$) is no worse than that, that we would have obtained if no misspecification had occurred. Naturally, because of market completeness, we also find in this case that the initial investment is as well greater than the one we would have had made, had we no uncertainty. If we would have assumed the opposite, that is, if we had an almost sure dominance of σ over $\bar{\sigma}$, the conclusions would be reversed. It is then by virtue of this theorem, that the robustness property of the Black-Scholes Model comes to light.

In section (4), we retain the main theorem of the previous section in mind, but now we approach a quite different problem. Contrary to the previous section where it is implicit that the true volatility $(\sigma(t,S_t))_{0\leqslant t\leqslant T}$ is fixed (assume that we fix a terminal time T), and additionally that it is of our interest to analyze cases where the misspecified volatility dominates (is dominated) by the latter; here we assume *a priori* a local reference volatility model given by $(\bar{\sigma}(t,S_t))_{0\leqslant t\leqslant T}$ governing the underlying asset, obtained by some numerical analysis on previous data. It is now of our interest to assess the performance of our hedging strategies, hence Definition (65) or P&L(Profit and Losses) and the concern now with optimality. We thus define the problem more concretely at hand and introduce a penalty function attached with an uncertainty parameter ψ. The uncertainty parameter measures how averse we are to the local reference model given by $(\bar{\sigma}(t,S_t))_{0\leqslant t\leqslant T}$ and the penalty function is identified with a function f satisfying some suitable properties. These two additions together, which turn the optimization problem into a penalized one, function as a measure of distance to our reference model identified by $(\bar{\sigma}(t,S_t))_{0\leqslant t\leqslant T}$, where f creates that measure of distance and ψ compounds that measure, so to speak. All in all, we arrive at the optimization problem at interest and define the value function u as in (72). More concretely, we now take a very abstract set consisting of natural (true) volatilities and hedging strategies and, according to optimization Problem (72), (and speaking very plainly) measure the performance of each hedging strategy "against"a true fixed volatility (which governs the asset), while holding to our model identified by $(\bar{\sigma}(t,S_t))_{t\geqslant 0}$. That is the reason why, in this setting, the true volatility may be coined as adversarial. Then, with a view towards an asymptotic explicit formula with regard to this optimal hedging strategy, in the sense discussed above (as well as to the correspondent price of the underlying option), we perform a Taylor expansion of the value function around $\psi = 0$ (where the uncertainity aversion is very small, meaning that we have a very good belief on our reference model) and arrive at Theorem (11). Here, one is faced with another function w (which is just the value function u viewed in another light), briefly introduced in (74) and described by Hypothesis (3), whose statements become at once comprehensible in subsection (4.4), in which an heuristic proof of Theorem (11) is provided. We then derive explicitly not only the full characterization of the value function w (thereby attaining the main result in (86)), but also the formulas for the optimal hedging strategy (83) and the adversarial volatilities (84).

Finally, in Section (5), we consider a novel multiplicative setting together with a logarithmic utility function, which is not covered under Hypothesis (4), in the spirit of the previous section. We begin by deriving the classical delta hedging strategy within this multiplicative setting, obtaining a formulation closely related to Lemma (8). We first analyze the case in which the true volatility process $(\sigma(t,S_t))_{t\geqslant 0}$ is fixed, allowing us to derive explicit expressions for the corresponding optimal trading strategies. Subsequently, we extend the analysis to a more general optimization problem by allowing the adversarial volatilities to vary freely within the admissible class. To address this setting, we introduce a penalty functional (more precisely, a quadratic error penalty) leading to the penalized optimization problem defined in (110). We then perform a Taylor expansion around $\psi = 0$ under Conjecture (1). By virtue of the multiplicative structure of the modified asset process introduced in (104), together with the geometric analogue of the profit-and-loss process defined in (105), the optimization problem can be evaluated directly at the terminal time T, without requiring a further extension of the value function u through function w, already mentioned. After extensive calculations, we conclude our thesis, by deriving explicit formulas for both the optimal hedging strategies and the corresponding adversarial volatilities.

2 General Overview of Stochastic Differential Equations

2.1 Definitions and Initial Considerations

We will therefore restrict our attention, for the purposes of this thesis, to stochastic differential equations (hereafter SDEs) of the form

$$dX_t = b(t, X.)\, dt + \sigma(t, X.)\, dB_t, \tag{1}$$

where $(B_t)_{t\geqslant 0}$ is a Brownian motion defined on a filtered probability space

$$(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geqslant 0}, \mathbb{P}).$$

Such equations arise naturally in mathematical finance, where (1) may be interpreted as describing the dynamics of the price of a financial asset.

Equation (1) is understood in the integral sense, namely

$$X_t = \xi + \int_0^t b(s, X.)\, ds + \int_0^t \sigma(s, X.)\, dB_s, \qquad t \geqslant 0, \tag{2}$$

where ξ is an $\mathcal{F}_0$-measurable random variable. Since $(X_t)_{t\geqslant 0}$ is adapted to the filtration $(\mathcal{F}_t)_{t\geqslant 0}$, it is natural to require that the processes $b(t, X_t)$ and $\sigma(t, X_t)$ be predictable, ensuring that both integrals in (2) are well defined.

Definition 1 (Solution of an SDE). *[13, Definition 1.1] Let*

$$b : [0, \infty) \times \mathbb{R}^n \to \mathbb{R}^n, \qquad \sigma : [0, \infty) \times \mathbb{R}^n \to \mathbb{R}^{n\times d}$$

be jointly Borel measurable functions, and let μ be a probability measure on $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$.

A solution of the stochastic differential equation

$$dX_t = b(t, X.)\, dt + \sigma(t, X.)\, dB_t, \qquad t \geqslant 0, \tag{3}$$

with initial distribution μ is a pair (B, X) of continuous adapted processes defined on a filtered probability space

$$(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P}), \qquad \mathbb{F} = (\mathcal{F}_t)_{t\geqslant 0},$$

such that:

1. *$B = (B_t)_{t\geqslant 0}$ is a standard $(\mathbb{F}, \mathbb{P})$-Brownian motion with values in $\mathbb{R}^d$;*
2. *the initial value X_0 has distribution μ;*
3. *the integrability conditions required to define the stochastic integrals hold, i.e. for all $t \geqslant 0$, $i \in \{1, \ldots, n\}$, and $j \in \{1, \ldots, d\}$,*

$$\int_0^t |b_i(s, X.)|\, ds < \infty \quad a.s., \qquad \int_0^t \sigma_{ij}^2(s, X.)\, ds < \infty \quad a.s.;$$

4. for every $i \in \{1,\dots,n\}$, the i-th component of $X = (X_t^{(1)},\dots,X_t^{(n)})_{t\geqslant 0}^\top$ satisfies, up to indistinguishability,

$$X_t^{(i)} = X_0^{(i)} + \int_0^t b_i(s,X.)\,ds + \sum_{j=1}^d \int_0^t \sigma_{ij}(s,X.)\,dB_s^{(j)}, \qquad t \geqslant 0. \tag{4}$$

Before proceeding further, the notion of predictability, although seemingly natural, deserves some justification. In particular, it plays a crucial role in ensuring the well-posedness of stochastic integrals. To that purpose, the following example shows that relaxing the predictability condition on the integrand may lead to inconsistencies.

Example 1. Consider the process $M_t = N_t - t$, where $(N_t)_{t\geqslant 0}$ is a Poisson process of unit intensity defined on a filtered probability space $(\Omega,\mathcal{F},(\mathcal{F}_t)_{t\geqslant 0},\mathbb{P})$. We first show that $(M_t)_{t\geqslant 0}$ is a martingale.

For $0 \leqslant s \leqslant t$, using the independent increments of the Poisson process,

$$\begin{aligned}
\mathbb{E}[M_t \mid \mathcal{F}_s] &= \mathbb{E}[N_t - t \mid \mathcal{F}_s]\\
&= \mathbb{E}[N_s + (N_t - N_s) \mid \mathcal{F}_s] - t\\
&= N_s + \mathbb{E}[N_t - N_s] - t\\
&= N_s + (t-s) - t\\
&= N_s - s = M_s.
\end{aligned}$$

Thus, $(M_t)_{t\geqslant 0}$ is a martingale.

Now consider the stochastic integral

$$\int_0^t N_s\,dM_s.$$

Using the decomposition $N_s = N_{s-} + \Delta N_s$, we obtain

$$\int_0^t N_s\,dM_s = \int_0^t N_{s-}\,dM_s + \int_0^t \Delta N_s\,dM_s = \int_0^t N_{s-}\,dM_s + \sum_{0<s\leqslant t} (\Delta N_s)^2.$$

Since $(N_t)_{t\geqslant 0}$ is a unit Poisson process, its jumps satisfy $\Delta N_s \in \{0,1\}$, hence $(\Delta N_s)^2 = \Delta N_s$, and therefore

$$\int_0^t N_s\,dM_s = \int_0^t N_{s-}\,dM_s + N_t.$$

The process $(N_{s-})_{s\geqslant 0}$ is predictable and locally bounded, so the first term is a local martingale. However, the second term N_t is not a martingale, and hence the full integral fails to be a local martingale. Indeed $(N_s)_{s\geqslant 0}$ is not predictable, even though it is adapted and càdlàg.

We are now in a position to revisit equation (1) more carefully. In order to formalize the notion of predictability for the coefficients, we introduce the canonical path space and the class of predictable path functionals acting on it. For this, we follow [1].

Definition 2. *[1, Chapter 5, Section 2, 8.2][Canonical Path Space] We let W (or, when necessary, W^d) denote the Polish space (Path Space)*

$$W := C([0,\infty), \mathbb{R}^d)$$

of continuous paths in $\mathbb{R}^d$, endowed with the topology of locally uniform convergence. The associated Borel σ-algebra is denoted by $\mathcal{W} := \mathcal{B}(W)$.

We define the canonical process $X = (X_t)_{t\geqslant 0}$ on W by

$$X_t(w) := w(t), \qquad t \geqslant 0,\ w \in W.$$

For $t \geqslant 0$, let

$$\mathcal{W}_t := \sigma(X_s : 0 \leqslant s \leqslant t), \qquad \mathcal{W}_\infty := \sigma(X_s : s \geqslant 0),$$

and call $\{\mathcal{W}_t\}_{t\geqslant 0}$ the canonical filtration.

Remark 1. *Note that the σ-algebras $\mathcal{W}_t$ are not completed.*

Remark 2. *If X is a continuous stochastic process, then adaptadness and predictability coincide.*

Definition 3. *[1, Chapter 5, Section 2, Definitionn 8.3][Predictable path functionals] Let $\{\mathcal{W}_t\}_{t\geqslant 0}$ denote the canonical filtration on the path space W, given by*

$$\mathcal{W}_t := \sigma(X_s : 0 \leqslant s \leqslant t),$$

where $X_t(w) = w(t)$ is the coordinate process on W. Let $\mathcal{P}^{\mathcal{W}}$ be the predictable σ-algebra on $(0,\infty) \times W$ associated with $\{\mathcal{W}_t\}_{t\geqslant 0}$, that is, the smallest σ-algebra on $\mathbb{R}_+ \times W$ such that every $\{\mathcal{W}_t\}$-adapted left-continuous process is measurable. A $\mathcal{P}^{\mathcal{W}}$-measurable mapping $a : \mathbb{R}_+ \times W \to \mathbb{R}$ is called a predictable path functional.

Example 2. *[1, Chapter 5, Section 2, 8.5]* Let $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geqslant 0}, \mathbb{P})$ be a filtered probability space, and $\mathcal{P}^{\mathcal{F}}$ denote the predictable σ-algebra on $(0,\infty) \times \Omega$ associated with $\{\mathcal{F}_t\}_{t\geqslant 0}$. Further, let $X = (X_t)_{t\geqslant 0}$ be a continuous $\mathbb{R}^d$-valued $\{\mathcal{F}_t\}$-adapted process. For each $\omega \in \Omega$, define the sample path

$$X.(\omega) : t \longmapsto X_t(\omega),$$

so that the path map $X. : \Omega \to W$ sends ω to its realized path. Consider the map

$$\Phi : (0,\infty) \times \Omega \to (0,\infty) \times W, \qquad \Phi(t,\omega) := (t, X.(\omega)).$$

We claim that Φ is $\mathcal{P}^{\mathcal{F}}/\mathcal{P}^{\mathcal{W}}$-measurable. To see this, it suffices to check that stopping times pull back correctly. If $T : W \to \mathbb{R}_+$ is an $\{\mathcal{W}_t\}$-stopping time, then for each $t \geqslant 0$

$$\{\omega : (T \circ X.)(\omega) \leqslant t\} = X.^{-1}\big(\{w : T(w) \leqslant t\}\big) \in \mathcal{F}_t.$$

Thus $T \circ X.$ is an $\{\mathcal{F}_t\}$-stopping time. It follows that Φ is $\mathcal{P}^{\mathcal{F}}/\mathcal{P}^{\mathcal{W}}$-measurable.

Consequently, if $a : \mathbb{R}_+ \times W \to \mathbb{R}$ is a predictable path functional, then the composed process

$$(t,\omega) \longmapsto a\big(t, X.(\omega)\big)$$

is $\mathcal{P}^{\mathcal{F}}$-measurable, i.e. predictable.

Remark 3. *In particular, if* $a : \mathbb{R}_+ \times W \to \mathbb{R}$ *is a predictable path functional of the form*

$$a(t, w) = \sigma(t, w(t)),$$

for some Borel measurable function $\sigma : [0, \infty) \times \mathbb{R}^d \to \mathbb{R}$*, then the process*

$$(t, \omega) \longmapsto \sigma(t, X.(\omega))$$

is predictable with respect to $(\mathcal{F}_t)_{t \geqslant 0}$*. The same holds for coefficients of the form* $b(t, X_.)$*.*

Now, in contrast with ordinary differential equations, a SDE is formulated relative to a filtered probability space equipped with a Brownian motion. Within this framework two notions arise: that of a *strong setting*, where the filtered probability space and the driving Brownian motion are given *a priori*; and that of a *weak setting*, where one only requires the existence of some filtered space and some Brownian motion for which the equation is satisfied. In the strong formulation, the emphasis is on constructing the solution process X_t explicitly on a prescribed stochastic basis, given the initial data and the driving noise. In the weak formulation, by contrast, the focus is on the distributional properties of the solution: both the filtered space and the Brownian motion are allowed to vary, provided that the resulting process satisfies the prescribed dynamics.

We will focus, in the present moment, on the strong solution formulation. In this setting, the stochastic basis and the driving Brownian motion are fixed in advance, and solutions are constructed as processes adapted to this prescribed noise. Consequently, it becomes meaningful to compare solutions path by path, that is, for each realization of the underlying randomness.

This observation leads naturally to the notion of pathwise uniqueness in the most immediate sense. Indeed, in the strong framework, uniqueness should reflect the idea that the driving Brownian motion completely determines the evolution of the system: if two solutions are defined on the same probability space, with the same Brownian motion and the same initial condition, then we expect they should coincide almost surely. By contrast, such a notion is not available in the weak formulation, where both the probability space and the Brownian motion are allowed to vary, and solutions are only specified in distribution. Thus, we introduce the following important definition, intimately associated with the strong formulation.

Definition 4. *[1, Chapter 5, Section 2, Definition 9.2][Pathwise uniqueness]*

We say that pathwise uniqueness *holds for* (3) *(with initial distribution* μ*) if the following statements are true:*

1. *For* any *filtered probability space*

   $$(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P}), \qquad \mathbb{F} = (\mathcal{F}_t)_{t \geqslant 0},$$

   satisfying the usual conditions and carrying a d-dimensional $(\mathbb{F}, \mathbb{P})$*-Brownian motion* $B = (B_t)_{t \geqslant 0}$*;*

2. *For* any *two* $\mathbb{F}$*-adapted continuous processes* $X = (X_t)_{t \geqslant 0}$ *and* $X' = (X'_t)_{t \geqslant 0}$ *(in particular, continuous semimartingales) such that:*

   (a) $X_0 = X'_0$ *a.s. and* $\mathcal{L}(X_0) = \mu$*;*

(b) *the integrability conditions required to define the Itô integrals hold, i.e. for all* $t \geqslant 0$ *and all components* $i \in \{1, \dots, n\}$, $j \in \{1, \dots, d\}$,

$$\int_0^t |b_i(s, X_\cdot)| \, ds < \infty, \qquad \int_0^t |\sigma_{ij}(s, X_\cdot)|^2 \, ds < \infty \quad a.s.,$$

and likewise with X *replaced by* X'*;*

(c) X *and* X' *both solve* (3) *in* integral form *with respect to the* same *Brownian motion* B*, i.e. for every* $i \in \{1, \dots, n\}$,

$$X_t^{(i)} = X_0^{(i)} + \int_0^t b_i(s, X_\cdot) \, ds + \sum_{j=1}^d \int_0^t \sigma_{ij}(s, X_\cdot) \, dB_s^{(j)},$$

$$X_t'^{(i)} = X_0'^{(i)} + \int_0^t b_i(s, X_\cdot') \, ds + \sum_{j=1}^d \int_0^t \sigma_{ij}(s, X_\cdot') \, dB_s^{(j)}, \qquad t \geqslant 0;$$

then

$$\mathbb{P}\big(X_t = X_t' \text{ for all } t \geqslant 0\big) = 1. \tag{5}$$

i.e. X *and* X' *are indistinguishable or, more concretely, there exists a set*

$$N \in \sigma\left(\bigcup_{t \geqslant 0} \mathcal{F}_t\right) \quad \text{with} \quad \mathbb{P}(N) = 0$$

such that

$$\{X_t \neq X_t' \text{ for some } t \in [0, \infty)\} \subseteq N$$

Our decision not to include the initial measure in the usual setup is motivated by the example below, which shows that pathwise uniqueness of solutions may depend heavily on the choice of the initial distribution. In particular, this phenomenon already appears in the case of a deterministic SDE.

Example 3. *[13, Example 1.5]* Fix $\alpha \in (0, 1)$ and consider the one–dimensional (deterministic) SDE

$$(X_t^+)^{-\alpha} \, dX_t = \frac{1}{1 - \alpha} \, dt, \qquad t \geqslant 0, \tag{6}$$

where $x^+ := \max\{x, 0\}$ denotes the positive part of $x \in \mathbb{R}$.

Assume first that the initial distribution X_0 satisfies

$$\mathcal{L}(X_0) = \mu \quad \text{and} \quad \mu\big((0, \infty)\big) = 1.$$

Then $X_0 > 0$ almost surely. Since the right–hand side of (6) is nonnegative, any solution has a nonnegative derivative whenever it is strictly positive. Consequently, for every $\omega \in \Omega$ such that $X_0(\omega) > 0$, we have

$$X_t(\omega) \geqslant X_0(\omega) > 0 \quad \text{for all } t \geqslant 0.$$

Thus $(X_t^+)^{-\alpha} = X_t^{-\alpha}$ for all $t \geqslant 0$, and (6) reduces pathwise to the ordinary differential equation

$$X_t^{-\alpha} \, dX_t = \frac{1}{1 - \alpha} \, dt. \tag{7}$$

Define $Y_t := X_t^{1-\alpha}$. Applying the chain rule yields

$$dY_t = (1-\alpha)X_t^{-\alpha}dX_t = dt,$$

and hence

$$X_t^{1-\alpha} = X_0^{1-\alpha} + t.$$

Therefore, the solution is given explicitly by

$$X_t = \left(X_0^{1-\alpha} + t\right)^{\frac{1}{1-\alpha}}, \qquad t \geqslant 0. \tag{8}$$

Since the drift coefficient $x \mapsto \frac{1}{1-\alpha}x^\alpha$ is locally Lipschitz on $(0, \infty)$ and the solution remains strictly positive, classical ordinary differential equation theory implies that (8) is the unique solution (up to indistinguishability). In particular, pathwise uniqueness holds for (6) with this choice of initial distribution.

Assume now that the initial distribution μ is the Dirac measure at 0, i.e. $\mu = \delta_0$. Let τ be an arbitrary stopping time with value in $[0, \infty)$. Define

$$X_t := (t-\tau)_+^{\frac{1}{1-\alpha}}, \qquad t \geqslant 0.$$

Then $X_t = 0$ for all $t \leqslant \tau$, while for $t > \tau$ the process follows the (unique) positive solution of (6) starting from 0 at time τ. Hence X_t is a solution of (6) with initial distribution δ_0.

Since τ is arbitrary, this yields different solutions, which clearly violates condition (5).

Once more, the relevance of Definition (4) is that, within a fixed stochastic basis $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ and for a given initial distribution μ, the driving Brownian motion determines the solution uniquely. More precisely, whenever pathwise uniqueness holds, there exists at most one adapted process satisfying (3) with the prescribed Brownian motion and initial condition. Hence, the following definition

Definition 5. *[1, Chapter 5, Section 2, Definition 9.4][Pathwise Exact SDE] We say that SDE* (2) *is a pathwise exact SDE, only if, given a setup* $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$*, with* $\mathbb{F} = (\mathcal{F}_t)_{t\geqslant 0}$*, there is at most a semi-martingale* X *(up to indistinguishability) such that conditions given by Definition* (1) *hold.*

In this sense, the strong formulation exhibits a deterministic character once the randomness has been fixed: the sample paths of the Brownian motion uniquely determine the sample paths of the solution.

2.2 Uniqueness and Existence Theorems

To establish the desired property of pathwise uniqueness (Definition 4), we first fix a suitable setup. Let

$$(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P}), \qquad \mathbb{F} = (\mathcal{F}_t)_{t\geqslant 0},$$

be a filtered probability space satisfying the usual conditions, and let

$$B = (B_t)_{t\geqslant 0}$$

be a standard $\mathbb{F}$-adapted Brownian motion. Let b and σ be predictable path functionals as in Definition 3, well-defined on the canonical path space of Definition 2, and, by Remark 3, predictable with respect to the filtration $\mathcal{F}_t$.

Moreover, let $\xi \in L^2(\Omega, \mathcal{F}_0, \mathbb{P})$. Then it turns out that equation (2) admits a unique solution provided the coefficients satisfy appropriate Lipschitz and integrability conditions. We now make these conditions precise.

Definition 6. *[1, Section 5, Chapter 2, 11.1][Lipschitz condition] Let $f : [0, \infty) \times \mathbb{R}^d \to \mathbb{R}^d$ be a predictable path functional. We say that f satisfies a Lipschitz condition if there exists a constant $K > 0$ such that for all $t \geqslant 0$ and all paths $w, \tilde{w} \in \mathbb{R}^d$,*

$$\|f(t, w) - f(t, \tilde{w})\| \leqslant K \sup_{0 \leqslant s \leqslant t} \|w(s) - \tilde{w}(s)\|.$$

Definition 7. *[1, Section 5, Chapter 2, 8.8][Integrability condition] Let $f, g : [0, \infty) \times \mathbb{R}^d \to \mathbb{R}^d$ be predictable path functionals as in (3). The pair (f, g) satisfies the integrability condition if*

$$\int_0^T \|f(s, X.)\|^2 + \|g(s, X.)\| \, ds < \infty \qquad a.s,$$

for every $T > 0$.

Remark 4. *In Definition 6, the norm $\| \cdot \|$ is the standard Euclidean norm on $\mathbb{R}^d$.*

Remark 5. *If there is some time T, for which the condition (7) is not satisfied, we can always localize, by defining a stopping time*

$$\tau_\infty := \inf \Big\{ t \geqslant 0 : \int_0^t \|f(s, X.)\|^2 + \|g(s, X.)\|^2 \, ds = \infty \Big\},$$

and then consider the solution only up to τ_∞. As a consequence, we can always assume that (7) holds.

Theorem 1. *[2, Section 6.1, Theorem 6.1] Suppose we have the general setting given by Definition (1) with the further property that $\xi \in L^2(\mathcal{F}_0)$ and the predictable functionals σ: $[0, \infty) \times \mathbb{R}^n \to \mathbb{R}^{n \times d}$ and b: $[0, \infty) \times \mathbb{R}^n \to \mathbb{R}^n$, satisfy conditions (6) and (7) accordingly. Moreover, suppose that for each $T > 0$, there exists a constant C_T for which*

$$|\sigma(s, w)| + |b(s, w)| \leqslant C_T(1 + w_T^*), \tag{9}$$

where $w_T^* := \sup_{s \leqslant T} |w_s|$, for all $s \leqslant T$. Then, equation (2) is exact as in (5).

The proof rests on 2 important results, for which the proofs will be provided.

Lemma 1. *[1, Chapter 5, Section 2, Lemma 11.5] Suppose we have a setup $(\Omega; \mathcal{F}; P; \mathcal{F}_t)$, an adapted process relative to the filtration $\mathcal{F}_t$ $X_t = (X_t^1, ..., X_t^n)$ and predictable path functionals $\alpha \in L^p(R^d, R^n)$ and $\beta \in L^p(R^n)$ defined as in (3). Suppose, that X satisfies equation (2) in matrix form, that is,*

$$X_t = \xi + \int_0^t \alpha(s, X.) \, ds + \int_0^t \beta(s, X.) \, dB_s, \qquad t \geqslant 0, \tag{10}$$

where $\xi \in L^p(\mathcal{F}_0)$ and B is a d-Brownian Motion.

Then, for each $T>0$ and $p \geqslant 2$, there exists a constant depending on t and p, $K_{t,p}$, such that

$$\mathbb{E}[(X_t^*)^p] \leqslant K_{T,p}\left(\mathbb{E}[|\xi|^p] + \mathbb{E}\left[\int_0^t \left(|\alpha_s|^p + |\beta_s|^p\right) ds\right]\right), \qquad 0 \leqslant t \leqslant T,$$

where $X_t^* := \sup_{0\leqslant s\leqslant t} |X_s|$.

Proof. First, note that

$$\begin{aligned}(X_t^*) &\leqslant |\xi| + \sup_{0\leqslant s\leqslant t}\left|\int_0^s \alpha_u \, dB_u\right| + \left(\int_0^t |\beta_u|, du\right) = A + B + C\\ &\implies (X_t^*)^p \leqslant (A+B+C)^p\end{aligned}$$

Now, consider a probability space $\Omega = \{1,2,3\}$ equipped with the uniform probability measure. Define a random variable Y by

$$Y(1) = A, \qquad Y(2) = B, \qquad Y(3) = C,$$

where A, B, C are defined above. By assumption $A, B, C \geqslant 0$ and since x^p is convex, Jensen´s inequality gives

$$\begin{aligned}&\left(\frac{A+B+C}{3}\right)^p \leqslant \frac{A^p + B^p + C^p}{3}\\ &\implies (X_t^*)^p \leqslant 3^{p-1}(A^p + B^p + C^p)\end{aligned}$$

Now, by Hölder´s inequality, we have

$$\begin{aligned}&\int_0^t |\beta_s| ds = \int_0^t |1 \cdot \beta_s| ds \leqslant \left(\int_0^t 1^{\frac{p}{p-1}} ds\right)^{\frac{p-1}{p}} \left(\int_0^t |\beta_s|^p ds\right)^{\frac{1}{p}} = t^{\frac{p-1}{p}} \left(\int_0^t |\beta_s|^p ds\right)^{\frac{1}{p}}\\ &\implies \mathbb{E}\left[C^p\right] = \mathbb{E}\left[\left(\int_0^t |\beta_s| ds\right)^p\right] \leqslant t^{p-1}\mathbb{E}\left[\int_0^t |\beta_s|^p ds\right]\end{aligned}$$

As for B^p, we apply the Burkholder-Davis-Gundy inequality proved in [1, Chapter 4, 42.1], obtaining

$$\mathbb{E}[B^p] \leqslant K_p \mathbb{E}\left[\left(\int_0^t \alpha_s^2 ds\right)^{\frac{p}{2}}\right]$$

Now, consider a probability space $([0,t], \mathcal{B}([0,t]))$ equipped with the uniform(continuous) probability measure. Define a random variable Y by

$$Y(s) = \alpha_s^2, \ 0 \leqslant s \leqslant t$$

Then, on the space above, with respect to the function $x^{\frac{p}{2}}$, which is convex by assumption, Jensen´s inequality gives

$$\left(t \cdot \frac{1}{t}\int_0^t \alpha_s^2 \, ds\right)^{\frac{p}{2}} = \left(\frac{1}{t}\int_0^t t\alpha_s^2 \, ds\right)^{\frac{p}{2}} \leqslant \frac{1}{t} \cdot \int_0^t t^{\frac{p}{2}} \alpha_s^p \, ds = t^{\frac{p}{2}-1} \int_0^t \alpha^p \, ds$$

Note that we are in the *strong* setting, meaning that we are given a realized Brownian path, i.e, a path is already fixed and therefore the above inequality only makes sense if we assume it to be pathwise. As a consequence, we obtain

$$\mathbb{E}\left[\left(\int_0^t \alpha_s^2 ds\right)^{\frac{p}{2}}\right] \leqslant t^{\frac{p}{2}-1}\mathbb{E}\left[\int_0^t \alpha_s^p ds\right]$$

Putting everything together, we have the result.

□

Corollary 1. *[1, Chapter 5, Section 2, Corollary 11.7] Suppose that* $p \geqslant 2$*,* $\xi_1, \xi_2 \in L^p(\mathcal{F}_0)$ *and let as usual* $b \in \mathbb{R}^n$ *and* $\sigma \in \mathbb{R}^{n\times d}$ *be predictable path functionals both satisfying the Lipschitz Condition* (6)*. Let* $X = (X_\cdot^1, ...X_\cdot^n)$ *and* $Y = (Y_\cdot^1, ...Y_\cdot^n)$ *be continuous adapted stochastic processes carried by some setup* $(\Omega, \mathcal{F}_{t\geqslant 0}, P)$*. Define two processes* $\tilde{X}$ *and* $\tilde{Y}$ *as*

$$\tilde{X}_\cdot = \xi_1 \ + \int_0^\cdot \alpha(s, X_\cdot)dB_s + \int_0^\cdot \beta(s, X_\cdot)ds$$

$$\tilde{Y}_\cdot = \xi_2 \ + \int_0^\cdot \alpha(s, Y_\cdot)dB_s + \int_0^\cdot \beta(s, Y_\cdot)ds$$

where B_s *is a Brownian motion on* $\mathbb{R}^d$*.*

Then, there exists a constant K_1 *depending on* t, n, p *and* K*, from* (1)*, such that*

$$\mathbb{E}[(\tilde{X} - \tilde{Y})_\cdot^*)^p] \leqslant K_1\left[\mathbb{E}[(|\xi_1 - \xi_2]|)^p] + \mathbb{E}\left[\int_0^\cdot ((X - Y)_s^*)^p ds\right]\right]$$

Proof. First, define the adapted continuous process $Z = \tilde{X} - \tilde{Y}$, which has the form

$$Z_t = (\xi_1 - \xi_2) + \int_0^t (\alpha(s, X_\cdot) - \alpha(s, Y_\cdot))dBs + \int_0^t (\beta(s, X_\cdot) - \beta(s, Y_\cdot))ds$$

Then apply (...) to the latter, thereby obtaining

$$\mathbb{E}[Z_\cdot^*)^p] \leqslant K_1\left[\mathbb{E}[(|\xi_1 - \xi_2]|)^p] + \mathbb{E}\left[\int_0^\cdot (\alpha(s, X_\cdot) - \alpha(s, Y_\cdot))^p + (\beta(s, X_\cdot) - \beta(s, Y_\cdot))^p ds\right]\right]$$

Applying then the Lipschitz condition (6) to both differences on the right term and absorbing the constants yields the result. □

Lemma 2. *Let* $f : [0, T] \to [0, \infty)$*, be a non-negative, continuous function defined in* $[0, T]$*. Assume that*

$$f(t) \leqslant c(t) + k\int_0^t f(s)\, ds$$

for $k \geqslant 0$ *and some integrable* $c : [0, \infty) \to \mathbb{R}^1$*. Then*

$$f(t) \leqslant c(t) + k\int_0^t c(s)e^{k(t-s)}ds, \quad 0 \leqslant t \leqslant T \tag{11}$$

Proof. By assumption and provided that f is non-negative, we have

$$\begin{aligned}
f(t) &\leqslant c(t) + k\int_0^t f(s)\,ds \leqslant c(t) + k\int_0^t \Big(c(s) + k\int_0^s f(u)\,du\Big)\,ds \\
&= c(t) + k\int_0^t c(s)\,ds + k^2\int_{0<u<s<t} f(u)\,du\,ds \\
&= c(t) + k\int_0^t c(s)\,ds + k^2\int_0^t\int_u^t f(u)\,ds\,du \\
&= c(t) + k\int_0^t c(s)\,ds + k^2\int_0^t f(s)(t-s)\,ds \qquad (12)
\end{aligned}$$

By induction on j, we prove now that for $j \geqslant 1$, it holds

$$f(t) \leqslant c(t) + \sum_{i=1}^{j} k^i \int_0^t c(s)\frac{(t-s)^{i-1}}{(i-1)!}ds \;+ k^{j+1}\int_0^t f(s)\frac{(t-s)^j}{j!}ds = g(t,j)$$

for $0 \leqslant t \leqslant T$. Indeed, for $j = 1$, $g(t,1)$ agrees with (12).

Fix t and suppose that $f(t) \leqslant g(t,j)$. Then by assumption if follows

$$\begin{aligned}
f(t) &\leqslant c(t) + \sum_{i=1}^{j} k^i \int_0^t c(s)\frac{(t-s)^{i-1}}{(i-1)!}ds \;+ k^{j+1}\int_0^t f(s)\frac{(t-s)^j}{j!}ds \\
&\leqslant c(t) + \sum_{i=1}^{j} k^i \int_0^t c(s)\frac{(t-s)^{i-1}}{(i-1)!}ds \;+ k^{j+1}\int_0^t (c(s) + k\int_0^s f(u)du)\frac{(t-s)^j}{j!}ds \\
&= c(t) + \sum_{i=1}^{j} k^i \int_0^t c(s)\frac{(t-s)^{i-1}}{(i-1)!}ds \;+ k^{j+1}\int_0^t c(s)\frac{(t-s)^j}{j!}ds \;+ k^{j+2}\int_{0<u<s<t}\frac{(t-s)^j}{j!}f(u)duds \\
&= c(t) + \sum_{i=1}^{j+1} k^i \int_0^t c(s)\frac{(t-s)^{i-1}}{(i-1)!}ds \;+ k^{j+2}\int_0^t f(u)\left[\int_u^t \frac{(t-s)^j}{j!}ds\right]du \\
&= c(t) + \sum_{i=1}^{j+1} k^i \int_0^t c(s)\frac{(t-s)^{i-1}}{(i-1)!}ds \;+ k^{j+2}\int_0^t f(u)\frac{(t-u)^{j+1}}{(j+1)!}du = g(t,j+1)
\end{aligned}$$

Since f is continuous and is defined on a compact set, it has a maximum M. Since, moreover, k is a fixed constant and $t - s \leqslant T$, then $\forall\ \epsilon > 0$, $\exists j_0 \geqslant 1$ such that $\forall j \geqslant j_0$, we have

$$k^{j+2}\int_0^t f(u)\frac{(t-u)^{j+1}}{(j+1)!}du \leqslant T^{j+1}Mk^{j+2}\int_0^t \frac{1}{(j+1)!}ds \leqslant \epsilon$$

If we let $\epsilon \to 0$, then $j \to \infty$ and since c is integrable on $[0,T]$ by assumption, it follows from the dominated convergence theorem that

$$\begin{aligned}
&f(t) \leqslant c(t) + \sum_{i=1}^{\infty} k^i \int_0^t c(s)\frac{(t-s)^{i-1}}{(i-1)!}ds = c(t) + k\int_0^t c(s)\sum_{i=1}^{\infty}\frac{(k(t-s))^{i-1}}{(i-1)!}ds \\
&= c(t) + k\int_0^t c(s)e^{k(t-s)}ds
\end{aligned}$$

□

Corollary 2. *[1, Chapter 5, Section 2, Exercise 11.11][Grönwall´s Lemma] Assume that all conditions in Lemma* (2) *hold. Assume, in addition, that* $c(t) = c$*, i.e, a constant on* $[0, T]$*. Then it follows that*

$$f(t) \leqslant ce^{kt}$$

Proof. It follows directly from solving (11), with $c(t) = c$. □

Proof of Theorem 1. We start by showing pathwise uniqueness. Fix a filtered probability space

$$(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P}), \qquad \mathbb{F} = (\mathcal{F}_t)_{t \geqslant 0},$$

satisfying the usual conditions, and let X and $\tilde{X}$ be two $\mathbb{F}$-adapted solutions of (2) with the same initial condition $\xi \in L^2(\mathcal{F}_0)$.

$$X_t = \xi + \int_0^t b(s, X.)\, ds + \int_0^t \sigma(s, X.)\, dB_s$$
$$\tilde{X}_t = \xi + \int_0^t b(s, \tilde{X}.)\, ds + \int_0^t \sigma(s, \tilde{X}.)\, dB_s$$

for $t \geqslant 0$. Notice that we are working implicitly with the augmented filtration

$$\mathbb{F}^{\mathbb{P}} := (\mathcal{F}_t^{\mathbb{P}})_{t \geqslant 0},$$

generated by the Brownian motion B and the initial sigma-algebra $\mathcal{F}_0$, and completed with respect to $\mathbb{P}$. In particular, all $\mathbb{P}$-null sets are included, which ensures that the solutions X and $\tilde{X}$ are $\mathbb{F}^{\mathbb{P}}$-adapted. Now, define a stopping time τ_m as

$$\tau_m := \inf\{t \geqslant 0 : \ |X_t - \tilde{X}_t| \geqslant m\}.$$

For fixed m, consider their correspondingly restriction to τ_m, that is

$$(X)_{t \wedge \tau_m} = \xi + \int_0^{t \wedge \tau_m} b(s, X.)\, ds + \int_0^{t \wedge \tau_m} \sigma(s, X.)\, dBs$$

$$(\tilde{X})_{t \wedge \tau_m} = \xi + \int_0^{t \wedge \tau_m} b(s, \tilde{X}.)\, ds + \int_0^{t \wedge \tau_m} \sigma(s, \tilde{X}.)\, dBs$$

for all $t \geqslant 0$. Now, by Corollary (1) with $p = 2$, we immediately obtain for a given $T \geqslant 0$,

$$\mathbb{E}[(X - \tilde{X})^*_{t \wedge \tau_m})^2] \leqslant K_1 \mathbb{E}\left[\int_0^{t \wedge \tau_m} ((X - \tilde{X})^*_s)^2 ds\right] = K_1 \int_0^{t \wedge \tau_m} \mathbb{E}\left[((X - \tilde{X})^*_s)^2\right] ds, \qquad 0 \leqslant t \leqslant T.$$

for a determined constant K_1, by Fubini´s Theorem. Define

$$f(t) = \mathbb{E}[(\tilde{X} - \tilde{X})^*_{t \wedge \tau_m})^2], \qquad 0 \leqslant t \leqslant T.$$

By definition, f is non-negative and with the dominated convergence theorem it follows that f continuous on $[0, T]$. Therefore, we can apply Gronwald´s Lemma (2) and together with $\tau_m \to \infty$, $f = 0$ on $[0, T]$. That implies $X = \tilde{X}$ on $[0, T]$ *a.s* and since T is arbitrary $X = \tilde{X}$ *a.s.* As for existence, first of all, it is enough to prove that a solution X exists on $[0, T]$. Indeed, assuming only existence on $[0, T]$, we can extend existence to (T, ∞). To that, consider a global solution X as well as a local solution X^T and any predictable path functional a as defined in (3). Then

$$a(s, X) = a(s, X_s) = a(s, X^T_s) = a(s, X^T)$$

for all $s \leqslant T$. Since T is arbitrary the extension follows. Fix $T > 0$ and let $\mathcal{L}_T$ be the space of continuous functions adapted to $\mathcal{F}_t$, for which $\mathbb{E}[((X_T)^*)^2] < \infty$ (and say, for the well posedness of the space, that $X_t := X_T$ for $t > T$). We state without proof that this space is Banach, as shown in [2, Page 158]. On this space, denote $\|X_t\| := \sqrt{\mathbb{E}[X^*_t]^2}$. We proceed by applying Picard´s iteration algorithm. For that purpose, define the map $\mathcal{R} : \mathcal{L}_T \to \mathcal{L}_T$ and the iterations X^n_t as follows

$$\begin{aligned}
&X^n_t = \mathcal{R}(X^{n-1}_t) = \xi + \int_0^t b(s, X^{n-1}_\cdot)\, ds + \int_0^t \sigma(s, X^{n-1}_\cdot)\, dBs, \qquad 0 \leqslant t \leqslant T \\
&X^0_t = \xi
\end{aligned}$$

Let us remember that $\xi \in L^2(\mathcal{F}_0)$ by assumption. Fix m and consider processes X^m_t and X^{m-1}_t. By Corollary (1), we obtain

$$\begin{aligned}
\|X^m_T - X^{m-1}_T\|^2 &= \mathbb{E}[((X^m_T - X^{m-1}_T)^*)^2] \leqslant K_1 \int_0^T \mathbb{E}\left[((X^{m-1} - X^{m-2})^*_s)^2\right] ds \\
&\leqslant K_1^m \int_{0<s_{m-1}<\ldots<s_1}^T \mathbb{E}\left[((X^1 - X^0)^*_{s_{m-1}})^2\right] ds_{m-1}\ldots ds_1 \\
&\leqslant K_1^m \int_{0<s_{m-1}<\ldots<s_1}^T \mathbb{E}\left[\int_0^T |b(s_{m-1}, \xi)|^2 + |\sigma(s_{m-1}, \xi)|^2\right] ds_{m-1}\ldots ds_1 \\
&\leqslant K_1^m \int_{0<s_{m-1}<\ldots<s_1}^T \mathbb{E}\left[\int_0^T K_2^2(1+\xi)^2\right] ds_{m-1}\ldots ds_1 \\
&\leqslant K_1^m \int_{0<s_{m-1}<\ldots<s_1}^T K_2^2 T(1 + \mathbb{E}\left[\xi^2\right]) ds_{m-1}\ldots ds_1 \\
&= K_1^m \frac{T^{m+1}}{m!} K_2^2(1 + \mathbb{E}[\xi^2]) \qquad (13)
\end{aligned}$$

where we used condition (9). Since T is fixed and $\xi \in L^2(\mathcal{F}_0)$, $m!$ dominates the sum, with the result that X^m is a Cauchy sequence on $\mathcal{L}_T$. Therefore, assume $X^m \to X$, as $m \to \infty$ in $\mathcal{L}_T$. To see that there exists indeed a solution, we notice that X is a fixed point with respect to $\mathcal{R}$. Indeed, we have the following

$$\|X - \mathcal{R}(X)\| \leqslant \|X - X^{m-1}\| + \|X^{m-1} - \mathcal{R}(X^{m-1})\| + \|\mathcal{R}(X^{m-1}) - \mathcal{R}(X)\|$$

As $m \to \infty$, the first and second terms go to 0, by assumption and by (13) respectively, since $\mathcal{R}(X^{m-1}) = X^m$. As for the last term, it suffices to see that, by definition of the map $\mathcal{R}$, we have

$$\|\mathcal{R}(X^{(m-1)}) - \mathcal{R}(X)\|^2_T \;\leqslant\; C_{T,K} \int_0^T \mathbb{E}\Big[\big((X^{(m-1)} - X)^*_s\big)^2\Big]\, ds.$$

By the dominant convergence theorem, the right term goes to 0 and thus $\mathcal{R}(X) = X$, thereby X is a solution of (2) on $[0, T]$. □

We thus have obtained that whenever the coefficients are globally Lipschitz and satisfy condition (9) (hereafter Linear Growth condition), together with the integrability of ξ, a solution X of (2) exists and, moreover, is unique. This result is enough for our future endeavors and it deserves some explanation. Nonetheless, it is of our interest to state and prove an even stronger result, which replaces the global Lipschitz condition to a local one.

Theorem 2. *[1, Chapter 5, Section 2, Theorem 12.1] Assume all conditions of Theorem (1) are satisfied. Further, assume that both predictable functionals σ and b are locally Lipschitz. More concretely, for every $N > 0$, suppose there exists $K_N > 0$ such that for all $t \leqslant T$*

$$\begin{aligned} |\sigma(t,w) - \sigma(t,\tilde{w})| + |b(t,w) - b(t,\tilde{w})| \leqslant K_N \sup_{0 \leqslant s \leqslant t} |w(s) - \tilde{w}(s)|, \\ \textit{for all } t \geqslant 0 \textit{ and } w, \tilde{w} \in W \textit{ such that } w^* \vee \tilde{w}^* \leqslant N \end{aligned} \tag{14}$$

Then, there exists a unique solution X that satisfies SDE (2).

Proof. Fix N and define the hitting time $\theta_N(w) := \inf\{t \geqslant 0 : \ |w_t| \geqslant N\} \wedge N$ and $\pi_N(w(t)) := w(t \wedge \theta_N(w))$. Define $\sigma_N(t,w) := \sigma(t, \pi_N w)$ and $b_N(t,w) := b(t, \pi_N w)$. Thus, condition (14) holds for σ_N and b_N for some K_N. Within these conditions, both coefficients are globally Lipschitz and thus if we let X_t^N be the process that satisfies

$$X_t^N = \xi + \int_0^t b_N(s, X_\cdot^N)\, ds + \int_0^t \sigma_N(s, X_\cdot^N)\, dBs, \qquad t \geqslant 0$$

by Theorem (1), X_t^N is thus the unique solution on $(0, \infty)$. We now construct a global solution by localization. For each $N \in \mathbb{N}$, define the stopping time

$$\tau_N := \inf\{t \geqslant 0 : |X_t^N| \geqslant N\} \wedge N.$$

By Corollary (1), it is clear that the processes X^{N+1} and X^N agree on $[0, \tau_N]$.

Hence, we may define a process X by setting

$$X_t := X_t^N \quad \text{for any } N \text{ such that } t < \tau_N.$$

Next, define

$$\tau_\infty := \sup_N \tau_N.$$

Then X is well-defined on $[0, \tau_\infty)$, and it remains to show that

$$\mathbb{P}(\tau_\infty = \infty) = 1,$$

so that X extends to a global solution on $[0, \infty)$.

To this end, fix $T > 0$. By the linear growth condition (9) and Lemma (1) (cf. the argument in (13)), there exists a constant $C > 0$ such that

$$f(t) = \mathbb{E}[((X_{t\wedge\tau_N})^*)^2] \leqslant C\left[1 + \mathbb{E}\left[\int_0^{t\wedge\tau_N} ((X_s)^*)^2 ds\right]\right]$$

Using Gröwnwall´s Lemma (2), we conclude that

$$f(t) \leqslant Ce^{Ct}$$

We are then able to show that $\tau_N^N = \infty$ $a.s$ using Markov inequality as follows

$$\begin{aligned}\mathbb{P}(\tau_N < T) &\leqslant \mathbb{P}\Big(\sup_{t\leqslant T} |X_{t\wedge\tau_N}| \geqslant N\Big)\\ &= \mathbb{P}\Big((X_{T\wedge\tau_N})^* \geqslant N\Big)\\ &= \mathbb{P}\Big(\big((X_{T\wedge\tau_N})^*\big)^2 \geqslant N^2\Big)\\ &\leqslant \frac{1}{N^2}\,\mathbb{E}\Big[\big((X_{T\wedge\tau_N})^*\big)^2\Big]\\ &\leqslant \frac{Ce^{CT}}{N^2} \longrightarrow 0, \quad \text{as } N \to \infty. \end{aligned} \tag{15}$$

To prove uniqueness, let X and Y be two solutions of the SDE and define this time

$$\tau_N := \inf\{t \geqslant 0 : |X_t| \vee |Y_t| \geqslant N\} \wedge N$$

Again, by Corollary (1), for some constant K_N, they must agree on $[0, \tau_N]$. Provided that $\sup_N \tau_N = \infty$ almost surely, we conclude $X = Y$ almost surely. To that, observe that for any $T > 0$,

$$\{\tau_N < T\} = \Big\{\sup_{t\leqslant T}(|X_t| \vee |Y_t|) \geqslant N\Big\} \subset \Big\{\sup_{t\leqslant T}|X_t| \geqslant N\Big\} \cup \Big\{\sup_{t\leqslant T}|Y_t| \geqslant N\Big\}.$$

Thus, we have

$$\mathbb{P}(\tau_N < T) \leqslant \mathbb{P}\Big(\sup_{t\leqslant T}|X_t| \geqslant N\Big) + \mathbb{P}\Big(\sup_{t\leqslant T}|Y_t| \geqslant N\Big).$$

which, by the estimates obtained above in (15), converges to 0 as $N \to \infty$, and hence

$$\mathbb{P}(\tau_N < T) \to 0.$$

Since $T > 0$ is arbitrary, it follows that $\sup_N \tau_N = \infty$ almost surely. □

On seeing the proofs more carefully, on the one hand it appears that a form of Lipschitz condition, either global or local, is necessary for uniquenesss to take place. On the other hand, it is not so clear whether the Lipschitz condition is necessary for the existence of a solution, since for the local version (2), it did not have as strong as a role as in the case of the global theorem (1) and, as remarked in [2, Page 159], the fixed point argument might not be the only way of proving existence. In addition to this, it is not so obvious whether we can obtain exactness, if we drop the Growth Condition (9) or if we just assume $\mathcal{F}_0$-measurability with respect to ξ. Both of these questions will remain unanswered in the present thesis. Moreover, as observed in [1, Chapter 5, Section 3, 16], both theorems provide conclusions that are much stronger than required, as they offer us a strong solution on any conceived setup (assuming square-integrability of ξ), which is not necessary for control problems. Thus, the concept of weak solution is also justified and deserves further, though small, clarification. Before turning to this, we next develop a deeper characterization of strong solutions, under which we will state and prove the diffeomoprhism theorem, a central result under the strong formulation.

2.3 Characterization of Strong Solutions

In this short subsection, we explore more openly the characterization of strong solutions and then connect it with the already well-established property of path-wise exactness of an SDE provided in (5). A strong solution usually means that, given a stochastic basis and Brownian motion, the solution is adapted to the filtration generated by the Brownian motion and the initial condition (which is already sufficient when, in the next section, we proceed with the proof the diffeomorphism theorem, a result within the strong formulation). However, we decided to explore the definition provided by [1], who gives an even stronger (and more rigorous) formulation than the probabilistic one described above. In trying to connect path-wise uniqueness with strong solutions more emphatically, he lets the latter be represented by a measurable map $F(\xi, B)$, which serves to unquestionably prove the deterministic nature of a strong solution, once the Brownian motion and the initial condition are fixed. Thus, we start with the following definition.

Definition 8. *[1, Chapter 5, Section 2, 10.9][Strong solution] A measurable function*

$$F : \mathbb{R}^n \times W^d \to W^n$$

is called a strong solution of the SDE (1) *if:*

(i) For every $t \geqslant 0$,

$$F^{-1}(\mathcal{W}_t^n) \subset \mathcal{B}(\mathbb{R}^n) \otimes \overline{\mathcal{W}}_t^d.$$

(ii) For every stochastic basis $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P}, \xi, B)$ *with Brownian motion* B*, the process*

$$X := F(\xi, B)$$

is an $(\mathcal{F}_t)$*-adapted solution of the SDE.*

With this definition, one is led to think intuitively that there should exist a measurable deterministic functional F_μ (depending on the initial law μ) such that $X = F_\mu(\xi, B)$ a.s.. The next theorem, for which the proof is provided in [1, Chapter 5, Section 2, Theorem 10.4], makes this notion precise.

Theorem 3. *Fix a stochastic basis* $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P}, \xi, B)$*. Suppose* X *is an* $(\mathcal{F}_t)$*-adapted solution of the SDE with initial law* μ*. Then there exists a measurable map*

$$F_\mu : \mathbb{R}^n \times W^d \to W^n$$

such that:

(i) For every $t \geqslant 0$*,* F_μ *is* $\mathcal{F}_t$ *-measurable*

(ii) $\mathbb{P}\big(F_\mu(\xi, B) = X\big) = 1$

Moreover, for any other stochastic basis $(\tilde{\Omega}, \tilde{\mathcal{F}}, (\tilde{\mathcal{F}}_t), \tilde{\mathbb{P}}, \tilde{\xi}, \tilde{B})$ *with* $\tilde{\xi}$ *law* μ*, the process* $\tilde{X} = F_\mu(\tilde{\xi}, \tilde{B})$ *a.s, satisfies*

$$\tilde{X}_t = \tilde{\xi} + \int_0^t b(s, \tilde{X}_\cdot)\, ds + \int_0^t \sigma(s, \tilde{X}_\cdot)\, dB_s, \qquad t \geqslant 0,$$

With this result, it is clear that under Lipschitz conditions on the coefficients given by Theorem (1), with which we obtain uniqueness, and assuming that Theorem (3) holds for every law μ, one then obtains an equivalence relation between strong solutions and pathwise exactness as in Definition (5). This generalization, however, is non-trivial. With the following theorem, one shows that indeed such strong solution exists.

Theorem 4. *[1, Chapter 5, Section 2, Theorem 13.1] Suppose we have the process X_t that satisfies:*

$$dX_t = \sigma(t, X_\cdot)dB_t + b(t, X_\cdot)dt$$

and that b and σ satisfy the conditions given in Theorem (1). Then, there exists a unique strong solution $F : \mathbb{R}^n \times W^d \to W^n$ such that for each $w \in \mathcal{W}^d$, $x \to F(x, w)$ is continuous from $\mathbb{R}^n \to \mathcal{W}^n$.

First we define the dyadic rationals and provide a proof of Kolmogorov´s Continuity Theorem, which will be essential for future results.

Definition 9 (Dyadic Rationals)**.** *Denote*

$$\mathbb{D} = \{j2^{-k},\ j, k \in \mathbb{Z}\}$$

the set of dyadic rationals on $\mathbb{R}$ and $\mathbb{D}^n = \mathbb{D} \times ... \times \mathbb{D}$ the set of dyadic rationals on $\mathbb{R}^n$.

Definition 10 (Hölder continuous function)**.** *Let $D \subset \mathbb{R}$ and let $f : D \to \mathbb{R}^d$ be a function. For $\gamma \in (0, 1]$, the function f is called Hölder continuous with exponent γ if there exists a constant $C > 0$ such that for all $x, y \in D$,*

$$|f(x) - f(y)| \leqslant C|x - y|^\gamma.$$

Theorem 5 (Kolmogorov Continuity Theorem)**.** *Let $(\Omega, \mathcal{F}, P)$ be a probability space and let $(X_t)_{t\in[0,T]}$ be a stochastic process with values in $\mathbb{R}^d$. Suppose there exist constants $\alpha > 0$, $\varepsilon > 0$ and $C > 0$ such that for all $s, t \in [0, T]$*

$$\mathbb{E}[|X_t - X_s|^\alpha] \leqslant C|t - s|^{1+\varepsilon}.$$

Then, there exists a modification $(\tilde{X}_t)_{t\in[0,T]}$ of (X_t) with continuous sample paths which are locally Hölder continuous of any order γ satisfying

$$0 < \gamma < \frac{\varepsilon}{\alpha}.$$

More precisely, there exists positive random variables h and $\delta > 0$ (both depending on w) such that

$$P\left(\sup_{\substack{s,t\in[0,T] \\ 0<t-s<h(\omega)}} \frac{|\tilde{X}_t(\omega) - \tilde{X}_s(\omega)|}{|t - s|^\gamma} \leqslant \delta(w)\right) = 1.$$

Proof. For both the statement and full proof, we refer to [3, Theorem 2.10]. □

Proof of Theorem (4). Under the conditions of Theorem (1), consider the probability measure μ as the Dirac measure on some point $x \in \mathbb{D}^n$. Thus we have, together with Theorem (3), that $F_{\delta_x} = F^x = X^x$ a.s, for X the unique solution of SDE (1).

Consider, further, that we have a fixed setup $(\Omega, \mathcal{F}, \mathbb{P})$. For $T > 0$, $p = n + 1$, notice that for $x, y \in \mathbb{D}^n$, it holds by Corollary (1) the following

$$0 \leqslant \mathbb{E}[((X^x - X^y)^*_\cdot)^p] \leqslant K_1 \left[|x - y|^{n+1} + \int_0^\cdot \mathbb{E}[((X^x - X^y)^*_\cdot)^p] ds \right]$$

which by Grönwall´s Lemma (2) gives

$$\mathbb{E}[((X^x - X^y)^*_\cdot)] \leqslant K_1(|x - y|)^p e^{K_1 p}$$

For a fixed t, consider the family of processes $\{X^x_t \ : x \in \mathbb{D}^n, 0 \leqslant t \leqslant T\}$. According to Kolmogorov´s Continuity Theorem (5), with probability 1, there exits a continuous extension of the latter to a family of processes $\{X^x_t \ : x \in \mathbb{R}^n, 0 \leqslant t \leqslant T\}$. We can assume that if $N_T \in \Omega$ is a set for which no continuous extension is possible, then $Q^d(N_T) = 0$ and further, defining $N = \bigcup_T N_T$, then for $t \geqslant 0$,

$$F^x(w) = \begin{cases} \lim\limits_{\substack{y \to x \\ y \in \mathbb{D}^n}} F^y(w) = \lim\limits_{\substack{y \to x \\ y \in \mathbb{D}^n}} X^y(w), & (w \notin N) \\ 0, & (w \in N) \end{cases}$$

Suppose now that the setup above is also carried out by some initial condition ξ. Consider a family of measurable functions $\psi_k : \mathbb{R}^n \to \mathbb{D}^n$, such that $|\psi_k(x) - x| \leqslant 2^{-k}$. Define a stochastic process X^k that satisfies SDE with initial condition

$$X^k_0 = \psi(\xi)$$

Then, by Theorem (3), there exists a mapping F for which $F(\psi(\xi), B_\cdot) = X^k_\cdot$ *a.s.* By Corollary (1), X^k converges uniformly on $(0, T)$ to X, with initial condition $X_0 = \xi$. But, by definition of the continuity of the map F, we have

$$X^k \to F(\xi, B_\cdot)$$

thereby $X = F(\xi, B_\cdot)$ a.s. Such construction then shows that for each w, $x \to F(x, w)$ is continuous and $w \to F(x, w)$ is $\bar{\mathcal{W}}^d/\mathcal{W}^n$ measurable for each x. Thus, it follows that for each t, $x \to F(x, w)$ is $\mathcal{B}(\mathbb{R}^d) \times \bar{\mathcal{W}}^d/\mathcal{W}^n$ and so, joint measurability follows. □

We moreover add the following result, which is of importance to the subsection that follows.

Theorem 6. *If the SDE is of diffusion type so that:*

$$\sigma(t, X_t) = \sigma(X_t) \qquad b(t, X_t) = b(X_t)$$

and both coefficients satisfy conditions of Theorem (1)*, then the strong solution F has the following property. For all $u \geqslant 0$ and x*

$$\mathbb{P}\big(X^x_{u+s}(w) = F\big(X^x_u(w), \theta_u w(t)\big) = 1$$

where $(\theta_u w(t)) = w_{u+s}$. For each x, the process $\{X^x_t : t \geqslant 0\}$ is Markovian.

Proof. We both statement and proof, we refer to [1, Chapter 5, Section 2, Lemma 13.6] □

2.4 The Diffeomorphism Theorem

We are thus ready to prove the most important result of this section. We will follow the stronger version provided by [4] that assumes boundness on all derivatives of the coefficients σ and b and assume from here on that the SDE is of diffusion type, so that by Theorem (6), for each initial condition x, we consider the Markovian process $\{X_t^x : t \geqslant 0\}$.

Theorem 7 (Diffeomorphism Theorem, C^∞ version)**.** *[4, Chapter 5, Theorem 2.3] Let*

$$\begin{aligned} &\sigma : \mathbb{R}^n \to \mathbb{R}^n \times \mathbb{R}^d \\ &b : \mathbb{R}^n \to \mathbb{R}^n \end{aligned}$$

be given such that they are smooth functions (C^∞) and all derivatives of σ_k^i and b^i are bounded. Assume moreover that they satisfy the Linear Growth Condition (9)*. Let X_t be the solution to the SDE on matrix form*

$$\begin{aligned} dX_t^x &= \sigma(X_t^x)dB_t + b(X_t^x)dt \\ X_0^x &= x, \qquad\qquad \text{for all } t \geqslant 0 \end{aligned} \tag{16}$$

Then, there exists a modification of X_t^x such that, for almost every $w \in \Omega$, the map $x \to X_t^x$ is a C^∞-diffeomorphism of $\mathbb{R}^n$ for each $t \geqslant 0$.

For this purpose, we begin with a lemma that uses polygonal approximations of paths. Set

$$\phi_n(s) = k\,2^{-n}, \qquad \text{for } s \in [k\,2^{-n},\, (k+1)\,2^{-n}).$$

We start by proving the following lemma.

Lemma 3. *[4, Chapter 5, Lemma 2.1] Let*

$$\begin{aligned} &\alpha : \mathbb{R}^n \to \mathbb{R}^n \times \mathbb{R}^d \\ &\beta : \mathbb{R}^n \to \mathbb{R}^n \end{aligned}$$

Assume further that for every $N > 0$ and $T > 0$, there exists a constant K_N for which

$$|\alpha(x) - \alpha(y)| + |\beta(x) - \beta(y)| \leqslant K_N|x - y|$$

for every $x, y \in \mathbb{R}^n$, with $|x|, |y| \leqslant N$. Additionally, assume that α and β satisfy the Linear Growth Condition (9)*. Further, let $c(t)$ and $c_n(t)$, $n = 1, 2, ...$ be continuous adapted processes such that for some $p \geqslant 2$*

$$\sup_n \mathbb{E}\left[\sup_{0\leqslant t\leqslant T} |c_n(t)|^{p+1}\right] < \infty \quad \textit{and} \tag{17}$$

$$\mathbb{E}\left[\sup_{0\leqslant t\leqslant T} |c(t) - c_n(t)|^p\right] \to 0, \quad n \to \infty \tag{18}$$

Let Y_t and Y_t^n, $n = 1, 2, ...$ be continuous predictable $\mathbb{R}^n$-processes which in matrix form satisfy

$$\begin{aligned} Y(t) &= c(t) + \int_0^t \alpha(Y(s))dB_s + \int_0^t \beta(Y(s))ds, \quad \textit{and} \\ Y^n(t) &= c_n(t) + \int_0^t \alpha(Y^n(\phi(s)))dB_s + \int_0^t \beta(Y^n(\phi(s)))ds \end{aligned}$$

for all $t \geqslant 0$. Then, for $T > 0$, we have

$$\mathbb{E}\left[\sup_{0\leqslant t\leqslant T} |Y(t) - Y^n(t)|^p\right] \to 0, \quad n \to \infty$$

Proof. First of all, notice that by Fatou´s Lemma we have

$$\begin{aligned}\mathbb{E}\left[\sup_{0\leqslant t\leqslant T} |c(t)|^{p+1}\right] &\leqslant \liminf_n \mathbb{E}\left[\sup_{0\leqslant t\leqslant T} |c_n(t)|^{p+1}\right] \\ &\leqslant \limsup_n \mathbb{E}\left[\sup_{0\leqslant t\leqslant T} |c_n(t)|^{p+1}\right] < \infty \end{aligned} \tag{19}$$

and also for $T > 0$,

$$\mathbb{E}\left[\sup_{0\leqslant t\leqslant T} |c_n(t) - c_n(\phi(t))|^p\right] \leqslant C_1 \left[\mathbb{E}\left[\sup_{0\leqslant t\leqslant T} |c(t) - c_n(\phi(t))|^p\right] + \mathbb{E}\left[\sup_{0\leqslant t\leqslant T} |c(t) - c_n(t)|^p\right]\right]$$

By the continuity of the polygon paths and by conditions (17) and (18), both terms on the right side converge to 0 by the dominated convergence theorem and hence we obtain

$$\mathbb{E}\left[\sup_{0\leqslant t\leqslant T} |c_n(t) - c_n(\phi(t))|^p\right] \to 0, \quad n \to \infty$$

Now fix some $N > 0$ and define, as before, the stopping times τ_n^N and τ^N as

$$\begin{aligned}\tau^N &= \inf\{t : |Y_t| > N\} \quad \text{and} \\ \tau_n^N &= \inf\{t : |Y_t^n| > N\}\end{aligned}$$

and observe that for a fixed $T > 0$, we have

$$\begin{aligned}&\mathbb{E}\left[\sup_{0\leqslant t\leqslant T} |Y(t) - Y^n(t)|^p\right] = \mathbb{E}\left[\sup_{0\leqslant t\leqslant T} |Y(t) - Y^n(t)|^p \mathbf{1}_{\tau_n^N \wedge \tau^N \leqslant T}\right] + \mathbb{E}\left[\sup_{0\leqslant t\leqslant T} |Y(t) - Y^n(t)|^p \mathbf{1}_{\tau_n^N \wedge \tau^N > T}\right] \\
&\leqslant \mathbb{E}\left[\sup_{0\leqslant t\leqslant T\wedge\tau_n^N\wedge\tau^N} |Y(t) - Y^n(t)|^p\right] + \mathbb{E}\left[\sup_{0\leqslant t\leqslant T} (|Y(t)| + |Y^n(t)|)^p : \tau^N > T\right] + \\
&\mathbb{E}\left[\sup_{0\leqslant t\leqslant T} (|Y(t)| + |Y^n(t)|)^p : \tau_n^N > T\right] \\
&= \mathbb{E}\left[\sup_{0\leqslant t\leqslant T\wedge\tau_n^N\wedge\tau^N} |Y(t) - Y^n(t)|^p\right] + \mathbb{E}\left[\sup_{0\leqslant t\leqslant T} (|Y(t)| + |Y^n(t)|)^p : Y(t) \leqslant N\right] + \\
&\mathbb{E}\left[\sup_{0\leqslant t\leqslant T} (|Y(t)| + |Y^n(t)|)^p : Y_n(t) \leqslant N\right] \\
&\leqslant \mathbb{E}\left[\sup_{0\leqslant t\leqslant T\wedge\tau_n^N\wedge\tau^N} |Y(t) - Y^n(t)|^p\right] + 2\mathbb{E}\left[\sup_{0\leqslant t\leqslant T} (|Y(t)| + |Y^n(t)|)^p : Y(t) + Y^n(t) \leqslant N\right] \\
&\leqslant \mathbb{E}\left[\sup_{0\leqslant t\leqslant T\wedge\tau_n^N\wedge\tau^N} |Y(t) - Y^n(t)|^p\right] + \frac{2}{N}\mathbb{E}\left[\sup_{0\leqslant t\leqslant T} (|Y(t)| + |Y^n(t)|)^{p+1}\right]\end{aligned}$$

Thus, provided that we can prove that the first term goes to 0 as $n \to \infty$ and that

$$\mathbb{E}\left[\sup_{0\leqslant t\leqslant T}|Y(t)|^{p+1}\right] < \infty \tag{20}$$

$$\mathbb{E}\left[\sup_{0\leqslant t\leqslant T}|Y^n(t)|^{p+1}\right] < \infty \tag{21}$$

then, as $N \to \infty$, we obtain the result, because for a fixed p, one then has

$$\mathbb{E}\left[\sup_{0\leqslant t\leqslant T}(|Y(t)| + |Y^n(t)|)^{p+1}\right] \leqslant \mathbb{E}\left[\sup_{0\leqslant t\leqslant T}(\max(|Y(t)|, |Y^n(t)|) + \max(|Y(t)|, |Y^n(t)|))^{p+1}\right]$$

$$= \mathbb{E}\left[\sup_{0\leqslant t\leqslant T}(2\max(|Y(t)|, |Y^n(t)|))^{p+1}\right] = 2^{p+1}\mathbb{E}\left[\sup_{0\leqslant t\leqslant T}(\max(|Y(t)|, |Y^n(t)|))^{p+1}\right] < \infty$$

For the first term, on the compact set $[0, T \wedge \tau_n^N \wedge \tau^N]$, both Y and Y^n are Lipschitz and so, for a fixed T, we can apply a slight variation of Corollary (1), due to the definition of Y^n, obtaining

$$\begin{aligned}
&\mathbb{E}[\sup_{0\leqslant t\leqslant T\wedge\tau_n^N\wedge\tau^N}|Y(t) - Y^n(t)|^p] \\
&\leqslant K_1\left[\mathbb{E}[\sup_{0\leqslant t\leqslant T}|c(t) - c_n(t)|^p] + \int_0^{T\wedge\tau_n^N\wedge\tau^N}\mathbb{E}\left[\sup_{0\leqslant u\leqslant s}|Y(u) - Y^n(\phi_n(u))|^p\right]ds\right] \\
&\leqslant K_2\mathbb{E}[\sup_{0\leqslant t\leqslant T}|c(t) - c_n(t)|^p] + K_2\int_0^{T\wedge\tau_n^N\wedge\tau^N}\mathbb{E}\left[\sup_{0\leqslant u\leqslant s}|Y(u) - Y^n((u))|^p\right] \\
&+ K_2\int_0^{T\wedge\tau_n^N\wedge\tau^N}\mathbb{E}\left[\sup_{0\leqslant u\leqslant s}|Y^n(u) - Y^n(\phi_n(u))|^p\right]ds \\
&= K_2\mathbb{E}[\sup_{0\leqslant t\leqslant T}|c(t) - c_n(t)|^p] + K_2\int_0^{T\wedge\tau_n^N\wedge\tau^N}\mathbb{E}\left[\sup_{0\leqslant u\leqslant s}|Y^n(u) - Y^n(\phi_n(u))|^p\right]ds \\
&+ K_2\int_0^{T\wedge\tau_n^N\wedge\tau^N}\mathbb{E}\left[\sup_{0\leqslant u\leqslant s}|Y(u) - Y^n((u))|^p\right]ds \\
&= K_2\left[\epsilon(n) + \mathbb{E}\left[\sup_{0\leqslant u\leqslant s}|Y(u) - Y^n((u))|^p\right]ds\right], \\
&\text{where } \epsilon(n) = \mathbb{E}[\sup_{0\leqslant t\leqslant T}|c(t) - c_n(t)|^p] + \int_0^{T\wedge\tau_n^N\wedge\tau^N}\mathbb{E}\left[\sup_{0\leqslant u\leqslant s}|Y^n(u) - Y^n(\phi_n(u))|^p\right]ds
\end{aligned}$$

Therefore, by Grönwald´s Lemma (2),

$$\mathbb{E}[\sup_{0\leqslant t\leqslant T\wedge\tau_n^N\wedge\tau^N}|Y(t) - Y^n(t)|^p] \leqslant \epsilon(n)e^{K_2T}$$

As a consequence, provided that $\epsilon(n) \to 0$, as $n \to \infty$, we obtain

$$\mathbb{E}[\sup_{0\leqslant t\leqslant T\wedge\tau_n^N\wedge\tau^N}|Y(t) - Y^n(t)|^p] \to 0$$

But $\epsilon(n) \to 0$ follows by noting that (18) holds on the first term and on the second term, given the continuity of the polygon paths, by applying the dominated convergence theorem. Hence, it takes only to prove now that

$$\mathbb{E}\left[(\sup_{0\leqslant t\leqslant T}|Y(t)|^{p+1}\right] < \infty$$

To that, observe that by definition

$$\mathbb{E}\left[\sup_{0\leqslant t\leqslant T}|Y(t)|^{p+1}\right]\leqslant C_1\mathbb{E}\left[\sup_{0\leqslant t\leqslant T}|c(t)|^{p+1}\right]+C_1\mathbb{E}\left[\sup_{0\leqslant t\leqslant T}\left|\int_0^t \alpha(Y(u))du\right|^{p+1}\right]$$
$$+C_1\mathbb{E}\left[\sup_{0\leqslant t\leqslant T}\left|\int_0^t \beta(Y(u))du\right|^{p+1}\right]$$

The first term is finite by (19). As for the second term, one has by application of the Burkholder-Davis-Gundy inequality proved in [1, Chapter 4, Section 6, 42.1], Hölder´s inequality and the Linear growth condition (9),

$$\mathbb{E}\left[\sup_{0\leqslant t\leqslant T}\left|\int_0^t \alpha(Y(u))du\right|^{p+1}\right]\leqslant C_2\mathbb{E}\left[\left|\int_0^t \alpha(Y(u))^2du\right|^{\frac{p+1}{2}}\right]$$
$$\leqslant C_3\int_0^t \mathbb{E}\left[|\alpha(Y(u))|^{p+1}\right]du\leqslant C_4\int_0^t \left[1+\mathbb{E}\left[|Y(u)|^{p+1}\right]\right]du$$

By a similar reasoning, one obtains

$$\mathbb{E}\left[\sup_{0\leqslant t\leqslant T}\left|\int_0^t \beta(Y(u))du\right|^{p+1}\right]\leqslant C_5\int_0^t \left[1+\mathbb{E}\left[|Y(u)|^{p+1}\right]\right]du$$

Hence, we have

$$\mathbb{E}\left[\sup_{0\leqslant t\leqslant T}|Y(t)|^{p+1}\right]\leqslant C_6\int_0^t \left[1+\mathbb{E}\left[|Y(u)|^{p+1}\right]\right]du$$

Thus, by Grönwald´s Lemma (2), we at last obtain

$$\mathbb{E}\left[\sup_{0\leqslant t\leqslant T}|Y(t)|^{p+1}\right]\leqslant C_7e^{C_7t}$$

Thus, we obtain (20). A similar reasoning can be obtained with regard to $\mathbb{E}\left[|Y^n(t)|^{p+1}\right]$. As a consequence, we have proven the result. □

This result already hints at what one ought to do next. We want to consider a solution of the SDE (16) and consider, for a fixed n, its transformation in the space of polygon paths. In the setting of the latter, by construction, we obtain directly nice regularity and measurability properties and, in particular, the so desired C^∞-map. Thus, we hope to extract a subsequence from it, such that it strongly approximates the original solution sufficiently well, or to be more precise, with respect to a specific norm that approximates the partial derivatives sufficiently well, and, at last, for almost all w, to obtain a modification of the original solution that satisfies the required conditions. To that end, fix a realization of the Brownian Motion B^α and consider the process X_t^n, which is defined on the canonical path space $(W,\mathcal{P}^W)$

$$X_t^n=x+\int_0^t\sigma(X_n(\phi(s))dB_s^\alpha+\int_0^t b(X_n(\phi(s))ds$$

for all $t\geqslant 0$. By definition of the polygon paths, for a fixed n and $t\in[(k-1)2^{-n},k2^{-n}]$, note that, given that $X_0^n=x$, one has

$$X_t^n=X_{(k-1)2^{-n}}^n+\sigma(X_{(k-1)2^{-n}}^n)\left[B_t^\alpha-B_{(k-1)2^{-n}}^\alpha\right]+b(X_{(k-1)2^{-n}}^n)\left[t-(k-1)2^{-n}\right]$$

Provided that both σ and b are C^{∞} functions, it is very clear that there exists some C^{∞} and measurable function F_n for which

$$X_t^n = F_n(x, B^{\alpha}_{\frac{1}{2^n}}, B^{\alpha}_{\frac{2}{2^n}}, ..., B^{\alpha}_{\frac{2^n-1}{2^n}}, B^{\alpha}_t)$$

for all $t \geqslant 0$. For a fixed d, define $\beta = (\beta_1, ..., \beta_d)$ such that

$$\beta = \beta_1 + ... + \beta_d$$

for $\beta_1, ..., \beta_d \in \mathbb{N}$ and a differential operator D_β such that

$$D_\beta = \frac{\partial x^\beta}{\partial x^{\beta_1} ... x^{\beta_d}} \tag{22}$$

and define $Y^{n,i}_{\beta,t} = D^\beta(X^{n,i}_t)$, where $1 \leqslant i \leqslant d$, represents each component of X_t^n. Then, we have the following Lemma

Lemma 4. *[4, Chapter 5, Proposition 2.1] For each $x \in \mathbb{R}^d$, $1 \leqslant i \leqslant d$ and each β, there exists a unique process such that*

$$\mathbb{E}\left[\sup_{0 \leqslant t \leqslant T} |Y^i_{\beta,t} - Y^{n,i}_{\beta,t}|^p\right] \to 0, \quad n \to \infty$$

for all $T > 0$ and $p \geqslant 1$. Moreover, we have uniform convergence on a bounded domain Ω, that is,

$$\sup_{x \in \Omega} \mathbb{E}\left[\sup_{0 \leqslant t \leqslant T} |Y^i_{\beta,t} - Y^{n,i}_{\beta,t}|^p\right] \to 0, \quad n \to \infty$$

Proof. Recall that, for $x \in \mathbb{R}^d$, we have

$$X_t^n = x + \int_0^t \sigma(X_n(\phi(s))dB_s + \int_0^t b(X_n(\phi(s))ds$$

Let us first consider $|\beta| = 1$. If we denote $Y^{n,i}_{t,j} = (Y^{n,i}_{t,j})_j$, where $\frac{\partial}{\partial x^j} X^{n,i}_t = Y^{n,i}_{t,j}$ then, deriving only once, we obtain in matrix form

$$Y_t^n = I + \int_0^t Y^n(\phi(s))\sigma'(X^n(\phi(s))dB_s + \int_0^t Y^n(\phi(s))b'(X^n(\phi(s))ds$$

where $b'(x)^i_j = \frac{\partial b^i}{\partial x^j}$ and $\sigma'(x)^i_j = \frac{\partial \sigma^i}{\partial x^j}$ denote the partial derivatives of b and σ for a fixed component i and naturally $I = \delta_{i,j}$. From Lemma (3), we obtain for $1 \leqslant i \leqslant d$

$$\mathbb{E}\left[\sup_{0 \leqslant t \leqslant T} |Y^i_t - Y^{n,i}_t|^p\right] \to 0, \quad n \to \infty$$

and

$$\mathbb{E}\left[\sup_{0 \leqslant t \leqslant T} |X^i_t - X^{n,i}_t|^p\right] \to 0, \quad n \to \infty$$

for each component i and for all $T > 0$, where Y^i_t is given by

$$Y^i_t = I + \int_0^t Y^i(s)\sigma'(X(s))dB_s + \int_0^t Y^i(s)b'(X(s))ds$$

Now, let us assume that $|\beta| = 2$, we are thus evaluating for second-order derivatives. In that case, set $\frac{\partial^2}{\partial x^i \partial x^j} X^{n,i}_{t,k,l} = Y^{n,i}_{t,k,l}$, $Y^{n,i}_{t,k,l} = (Y^{n,i}_{t,k,l})_{k,l}$ and we obtain

$$
\begin{aligned}
Y^{n,i}_{t,j_1,j_2} &= \int_0^t Y^{n,i}_{t,j_1,j_2}(\phi(s))\sigma'(X^n(\phi(s))dB_s + \int_0^t Y^{n,i}_{t,j_1,j_2} b'(X^n(\phi(s))ds \\
&+ \int_0^t \sigma''(X^n(\phi(s))Y^{n,i}_{t,j_1}(\phi(s))Y^{n,i}_{t,j_2}(\phi(s))dB_s \\
&+ \int_0^t b''(X^n(\phi(s))Y^{n,i}_{t,j_1}(\phi(s))Y^{n,i}_{t,j_2}(\phi(s))ds
\end{aligned}
$$

where $b''(x)^i_{j_1,j_2} = \frac{\partial^2 b^i}{\partial x^{j_1} \partial x^{j_2}}$ and $\sigma''(x)^i_{j_1,j_2} = \frac{\partial^2 \sigma^i}{\partial x^{j_1} \partial x^{j_2}}$. In addition, set

$$
\begin{aligned}
\alpha_n(t) &= \int_0^t \sigma''(X^n(\phi(s))Y^{n,i}_{t,j_1}(\phi(s))Y^{n,i}_{t,j_2}(\phi(s))dB_s \\
&+ \int_0^t b''(X^n(\phi(s))Y^{n,i}_{t,j_1}(\phi(s))Y^{n,i}_{t,j_2}(\phi(s))ds
\end{aligned}
$$

We can now apply Lemma (3) to

$$
Y^{n,i}_{t,j_1,j_2} = \alpha_n(t) + \int_0^t Y^{n,i}_{t,j_1,j_2}(\phi(s))\sigma'(X^n(\phi(s))dB_s + \int_0^t Y^{n,i}_{t,j_1,j_2} b'(X^n(\phi(s))ds
$$

obtaining, accordingly, for each $1 \leqslant i, j_1, j_2 \leqslant d$

$$
\mathbb{E}\left[\sup_{0 \leqslant t \leqslant T} |Y^i_{t,j_1,j_2} - Y^{n,i}_{t,j_1,j_2}|^p\right] \to 0, \quad n \to \infty
$$

where

$$
Y^i_{t,j_1,j_2} = \alpha(t) + \int_0^t Y^i_{t,j_1,j_2}(\phi(s))\sigma'(X(\phi(s))dB_s + \int_0^t Y^i_{t,j_1,j_2} b'(X(\phi(s))ds
$$

with $\alpha(t)$ correspondingly as

$$
\begin{aligned}
\alpha(t) &= \int_0^t \sigma''(X(s))Y^i_{t,j_1}(s)Y^i_{t,j_2}(s)dB_s \\
&+ \int_0^t b''(X(s))Y^i_{t,j_1}(s)Y^i_{t,j_2}(s)ds
\end{aligned}
$$

It is also clear that, on a compact set, uniform convergence holds. Thus, if we proceed with the same reasoning for higher-order derivatives, we obtain the result. □

Proposition 1. *[4, Chapter 5, Proposition 2.2][Stochastic Flow] Let X be the solution of the SDE* (16)*, whose coefficients satisfy conditions stated in* (7)*. Then, there exists a continuous modification of X, denoted by $\tilde{X}$, such that for each t, the mapping $x \in \mathbb{R}^d \to \tilde{X}(t, x, w)$ is C^∞ a.s.*

The above proposition is a fundamental result in the theory of stochastic differential equations, as it shows that the same C^∞ mapping (provided that it is also a diffeomorphism) generates, as the name suggests, a stochastic difeomorphic flow. In particular, for almost every ω, the mapping

$$
x \longmapsto \tilde{X}(t, x, \omega)
$$

is smooth (as well as its inverse) and bijective for each fixed $t \geqslant 0$, which is a much stronger property than mere existence and uniqueness of a solution. An important consequence of this regularity, coupled with global boundness on every derivative of the coefficients σ and b (though these conditions are too strong), is a strengthened version of the well-known property of non-crossing for different initial conditions. In other words, we obtain smooth trajectories which never intersect. The proof of this will be provided.

Proof of Proposition (1). Consider a smooth function f and fix $p \geqslant 1$. Recall the differential operator D_β defined on (22) and consider a bounded domain $\Omega \in \mathbb{R}^d$. Let $m \in \mathbb{N}$ and define the following semi-norms

$$\|f\|_{p,m}^{\Omega} = \sum_{|\beta| \leqslant m} \left[\int_\Omega |D_\beta f(x)|^p dx \right]^{\frac{1}{p}}, \quad \|f\|_{\infty,m}^{\Omega} = \sum_{\beta \leqslant m} \sup_{x \in \Omega} |D_\beta f(x)|$$

Now, it follows from the inequalities of Sobolev [5], that for $m' > m$, $\Omega \subset \Omega'$ and $p \geqslant 1$, there exists a constant C such that

$$\|f\|_{\infty,m}^{\Omega} \leqslant C \|f\|_{p,m'}^{\Omega'} \tag{23}$$

Thus, by use of previous Lemma (3), we know that for each β and m and for any compact domain Ω we must have component wise

$$\sup_{x \in \Omega} \mathbb{E} \left[\sup_{0 \leqslant t \leqslant T} |Y_{\beta,t}^{i} - Y_{\beta,t}^{n,i}|^p \right] \to 0, \quad n \to \infty$$

where

$$Y_{\beta,t}^{n,i} = D_\beta X_t^{n,i}, \quad Y_{\beta,t}^{i} = D_\beta X_t^{i}$$

for every $T > 0$. Thus, we have

$$\begin{aligned} &\mathbb{E} \left[\sup_{0 \leqslant t \leqslant T} \int_\Omega |Y_{\beta,t}^{i}(x) - Y_{\beta,t}^{n,i}(x)|^p dx \right] \\ &\leqslant \int_\Omega \mathbb{E} \left[\sup_{0 \leqslant t \leqslant T} |Y_{\beta,t}^{i}(x) - Y_{\beta,t}^{n,i}(x)|^p \right] dx \to 0, \quad n \to \infty \end{aligned}$$

We must have have then, by definition of $\|.\|_{p,m}^{\Omega}$, the following

$$\mathbb{E} \left[\sup_{0 \leqslant t \leqslant T} \|X_t^{n_1,i} - X_t^{n_2,i}\|_{p,m}^{\Omega} \right] \to 0$$

for all $n_1, n_2 \to \infty$, $T > 0$ and $p \geqslant 1$. Thus, there must exist some subsequence $X_t^{k,i}$, such that

$$\sup_{0 \leqslant t \leqslant T} \|X_t^{k_1,i} - X_t^{k_2,i}\|_{p,m}^{\Omega} \to 0$$

for all $k_1, k_2 \to \infty$. By (23), it follows that there must exist some subsequence, again $X_t^{n,i}$ such that

$$\sup_{0 \leqslant t \leqslant T} \|X_t^{k_1,i} - X_t^{k_2,i}\|_{\infty,m}^{\Omega} \to 0$$

for all $k_1, k_2 \to \infty$. That is, there exists a modification $\tilde{X}$, such that for almost all w,

$$\lim_{n \to \infty} X^n(t, x, w) = \tilde{X}(t, x, w).$$

exists uniformly for each (t,x) on each compact set in $[0,\infty]\times\mathbb{R}^d$. Also that it is continuous and for each $t\geqslant 0$, the map $x\to\tilde{X}(t,x,w)$ is C^∞ by construction. In particular, we obtain

$$P[X(t,x,w)=\tilde{X}(t,x,w)\,:\,t\geqslant 0]=1$$

for each x. □

We are left to show that it is indeed a diffeomorphism. We will first prove that the inverse mapping is C^∞.

Proposition 2. *Let X be a solution of the SDE* (16)*, whose coefficients satisfy the Linear Growth condition* (9)*. Assume, moreover, that there exists a modification of X, call it X again, such that for each $t\geqslant 0$, the continuous mapping $x\to X(t,x,w)$ is C^∞ a.s. Then, the inverse mapping $X(t,x,w)\to x$ is C^∞.*

Before proving this result, we will take a step back and take a look at the so called Stratonovich integrals. They will be important to simplify the proof of the above Proposition. To that end, let H_t and H'_t be two predictable stochastic processes adapted to some filtration $\mathcal{F}_t$. Then define

$$H_t\circ dH'_t:=H_tdH'_t+\frac{1}{2}d[H,H']_t$$

where $[.,.]$ denotes the co-variation between 2 stochastic processes. From the above representation it follows immediately by Itô that

$$\begin{aligned}d(H_tH'_t)&=H_tdH'_t+H'_tdH_t+d[H,H']_t\\&=H_tdH'_t+\frac{1}{2}d[H,H']_t+H'_tdH_t+\frac{1}{2}d[H,H']_t\\&=H_t\circ dH'_t+H'_t\circ dH_t\end{aligned}$$

After a small motivation, we enumerate several properties of the Stratonovich integrals in the following Proposition.

Proposition 3. *Let X and Y be continuous semimartingales. The Stratonovich integral*

$$\int_0^t X_s\circ dY_s$$

satisfies the following properties:

(i) ***Linearity.*** *For any constants $a,b\in\mathbb{R}$,*

$$\int_0^t(aX_s+bZ_s)\circ dY_s=a\int_0^t X_s\circ dY_s+b\int_0^t Z_s\circ dY_s.$$

(ii) ***Additivity in the integrator.***

$$\int_0^t X_s\circ d(Y_s+Z_s)=\int_0^t X_s\circ dY_s+\int_0^t X_s\circ dZ_s.$$

(iii) ***Classical chain rule.*** *If $f \in C^1(\mathbb{R})$, then*

$$f(X_t) = f(X_0) + \int_0^t f'(X_s) \circ dX_s.$$

(iv) ***Product rule***

$$X_t Y_t = X_0 Y_0 + \int_0^t X_s \circ dY_s + \int_0^t Y_s \circ dX_s.$$

(v) ***Relation to Itô integral.***

$$\int_0^t X_s \circ dY_s = \int_0^t X_s \, dY_s + \frac{1}{2}[X, Y]_t,$$

We will, from now on, assume all properties listed above. Assume that $H_t = \sigma(X_t)$ and $H'_t = B_t$ denote a Brownian Motion adapted to $\mathcal{F}_t$. For simplicity assume that $\sigma : \mathbb{R} \to \mathbb{R} \times \mathbb{R}$ and B_t lives on $\mathbb{R}^1$. Then, note by Itô that, and under SDE (16), we have

$$\begin{aligned} d(\sigma(X_t)) &= \sigma'(X_t) dX_t + \frac{1}{2}\sigma''(X_t) d[X, X]_t \\ &= \sigma'(X_t)\sigma(X_t) dB_t + \sigma'(X_t) b(X_t) dt + \frac{1}{2}\sigma''(X_t)\sigma(X_t)^2 dt \end{aligned}$$

and therefore,

$$\sigma(X_t) \circ dB_t = \sigma_t dB_t + \sigma'(X_t)\sigma(X_t) dt$$

And hence this particular representation gives rise to following equivalent expression of SDE (16)

$$dX_t = \sigma(X_t) \circ dB_t + \tilde{b}(X_t) dt,$$

where $\tilde{b}(X_t) = b(X_t) - \sigma'(X_t)\sigma(X_t)$.

Proof of Proposition (2). Assume that SDE (16) has representation

$$dX_t = \sigma(X_t) \circ dB_t + b(X_t) dt,$$

For simplicity, we can always assume that X_t is governed by this representation, since both drifts (the original one and the adjusted one) satisfy our requirements. For a true equivalence, a change of variables would be needed, but it would lead to the same conclusion, that of the existence of an inverse. Define, therefore, $Y_t^{i,j} = \frac{\partial}{\partial x^j} X_t^i$. We then have in matrix form the following

$$Y_t = I + \int_0^t \sigma'(X(s)) Y(s) \circ dB_s + \int_0^t Y(s) b'(X(s)) ds$$

Now, let $Z_t = (Z_t^{i,j})_{i,j}$ be the solution of

$$Z_t = I - \int_0^t \sigma'(X(s)) Z(s) \circ dB_s - \int_0^t Z(s) b'(X(s)) ds$$

Then, according to the previous remark, we have

$$
\begin{aligned}
d(Y_tZ_t) &= Y_t \circ dZ_t + Z_t \circ dY_t \\
&= Y_t \circ (-Z_t\sigma'(X_t) \circ dB_t - Z_tb'(X_t)dt) \\
&+ Z_t \circ (Y_t\sigma'(X_t) \circ dB_t + Y_tb'(X_t)dt) \\
&= -Y_tZ_t\sigma'(X_t) \circ dB_t - Y_tZ_tb'(X_t)dt \\
&+ Z_tY_t\sigma'(X_t) \circ dB_t + Z_tY_tb'(X_t)dt \\
&= 0
\end{aligned}
$$

Thus, we immediately obtain by integrating and $Y_0Z_0 = I$, that

$$Y_tZ_t = I$$

or, in other words, we have found that Y_t is invertible and, in particular, its inverse is Z_t. By construction, Z_t is C^∞, hence the result follows. ☐

Finally, we prove that the map is bijective.

Lemma 5. *[4, Chapter 5, Lemma 2.2] Let $\tilde{X}(t, x, w)$ be the process obtained in* (1) *which solves equation:*

$$
\begin{aligned}
&dX_t = \sigma(X_t) \circ dB_t - b(X_t)dt \\
&X(0) = x
\end{aligned}
$$

Then, we have

$$X(T - t, x, w) = \tilde{X}(t, X(T, x, w), \tilde{w}), \quad 0 \leqslant t \leqslant T$$

holds a.s, where $B(T - t) - B(T) = \tilde{B}(t)$, for all $0 \leqslant t \leqslant T$ is reversed Brownian path of B_t.

Proof. We will fix a discrete realization of the Brownian Motion B_n^α. Let then:

$$B_n^\alpha(t, w) = \frac{\bar{t}_n - t}{\bar{t}_n - t_n}\, w^\alpha(t_n) + \frac{t - t_n}{\bar{t}_n - t_n}\, w^\alpha(\bar{t}_n), \qquad \alpha = 1, 2, \dots, r,$$

where

$$\bar{t}_n = \frac{k+1}{2^n}T, \qquad t_n = \frac{k}{2^n}T \quad \text{if} \quad \frac{k}{2^n}T \leqslant t < \frac{k+1}{2^n}T,$$

$k = 1, 2, \dots, 2^n - 1$, $n = 1, 2, \dots$. It is clear that

$$B_n^\alpha(T - t, w) - w^\alpha(T) = B_n^\alpha(t, \hat{w}), \qquad 0 \leqslant t \leqslant T.$$

Both the following equations

$$
\begin{cases}
\dfrac{dX_n^i}{dt}(t) = \displaystyle\sum_{\alpha=1}^{r} \sigma_\alpha^i(X_n(t))\, \frac{dB_n^\alpha}{dt}(t, w) + b^i(X_n(t)), \quad i = 1, 2, \dots, d, \\
X_n(0) = x.
\end{cases}
$$

$$
\begin{cases}
\dfrac{dX_n^i}{dt}(t) = \displaystyle\sum_{\alpha=1}^{r} \sigma_\alpha^i(X_n(t))\, \frac{dB_n^\alpha}{dt}(t, w) - b^i(X_n(t)), \quad i = 1, 2, \dots, d, \\
X_n(0) = x.
\end{cases}
$$

have unique solutions X^n and $\tilde{X}^n$. It follows very easily by uniqueness that

$$X(T-t,x,w) = \tilde{X}(t, X(T,x,w), \tilde{w}), \quad 0 \leqslant t \leqslant T$$

$\square$

Corollary 3. *It follows easily that $x = X(t, \tilde{X}(t,x,\tilde{w}), w)$ and $x = \tilde{X}(t, X(t,x,w), \tilde{w})$, i.e, the continuous map $x \to X(t,x,w)$ is bijective.*

Remark 6. *Corollary (4) above essentially means that, by fixing $T > 0$, if we take a realization w under finite time horizon $[0,T]$ (that is, from 0 to T) and its reverse (that is, from T to 0) and let x be any point of the path X_t on $[0,T]$, that we can achieve this very same point x either by taking the path backwards for some time t and then going forward for t units of time again, $X(t, \tilde{X}(t,x,\tilde{w}), w)$; or by taking the path forward for some time t and then going backwards for t unitst of time again, $\tilde{X}(t, X(t,x,w), \tilde{w})$ (for an appropriate time t which remains under time horizon $[0,T]$), thus proving bijectivity.*

Corollary 4. *From Proposition (2) and Corollary (4), the Diffeomorphism Theorem (7) follows.*

The smoothness assumptions on the coefficients σ and b in Theorem 7 can, as one might expect, be weakened. If σ and b are only C^k (for some integer $k \geqslant 1$) with bounded derivatives up to order k, then the associated stochastic flow is no longer C^∞, but still retains C^k-regularity. We will state this theorem without proof.

Theorem 8 (Diffeomorphism Theorem, C^k-version)**.** *Let*

$$\sigma : \mathbb{R}^n \to \mathbb{R}^{n\times d}, \qquad b : \mathbb{R}^n \to \mathbb{R}^n$$

be C^k-functions for some integer $k \geqslant 1$, and assume that all derivatives of σ^i_k and b^i up to order k are bounded. Suppose moreover that σ and b satisfy the Linear Growth Condition (9).

Let X^x_t be the solution to the SDE

$$dX^x_t = \sigma(X^x_t)\, dB_t + b(X^x_t)\, dt, \tag{24}$$

$$X^x_0 = x, \qquad t \geqslant 0. \tag{25}$$

Then, for each ω, there exists a modification of X^x_t such that, for every $t \geqslant 0$, the map

$$x \mapsto X^x_t(\omega)$$

is a C^k-diffeomorphism of $\mathbb{R}^n$.

Proof. We refer to [8, Chapter 4.7, Theorem 4.7.2] $\square$

Proposition 4 (The Non-crossing Lemma: Strong Version)**.** *Let $X(t,x,\omega)$ be the stochastic flow generated by the SDE*

$$dX_t = \sigma(X_t)\, dB_t + b(X_t)\, dt,$$

where both b and σ are C^k functions in which all first k derivatives are globally bounded. Assume, moreover, that for each $t \geqslant 0$, the map

$$x \mapsto X(t,x,\omega)$$

is a C^k-diffeomorphism almost surely.

Then for any $x_1 < x_2$ we have

$$X(t, x_1, \omega) < X(t, x_2, \omega) \quad \textit{for all } t \geqslant 0$$

almost surely.

Proof. Fix $x_1 < x_2$ and define the hitting time

$$\tau := \inf\{t > 0 : X(t, x_1) = X(t, x_2)\}.$$

Suppose $\mathbb{P}(\tau < \infty) > 0$. If not, by the continuity of the paths, we have the result. Otherwise, by this very same continuity of the paths, at $t = \tau$, we therefore must have

$$X(\tau, x_1) = X(\tau, x_2).$$

Now, by assumption, both are, in particular, Markov processes. As a consequence, if they were to meet, they would become indistinguishable: from τ onward, they would remain the same *ad infinitum.* Thus, the weak version of the Non-Crossing Lemma follows. That is,

$$x_1 < x_2 \implies X(t, x_1, \omega) \leqslant X(t, x_2, \omega) \quad \forall t \geqslant 0. \tag{26}$$

In addition to this, by Corollary (4), for each path w

$$x \neq y \implies X(t, x, w) \neq X(t, y, w), \quad \forall t \geqslant 0$$

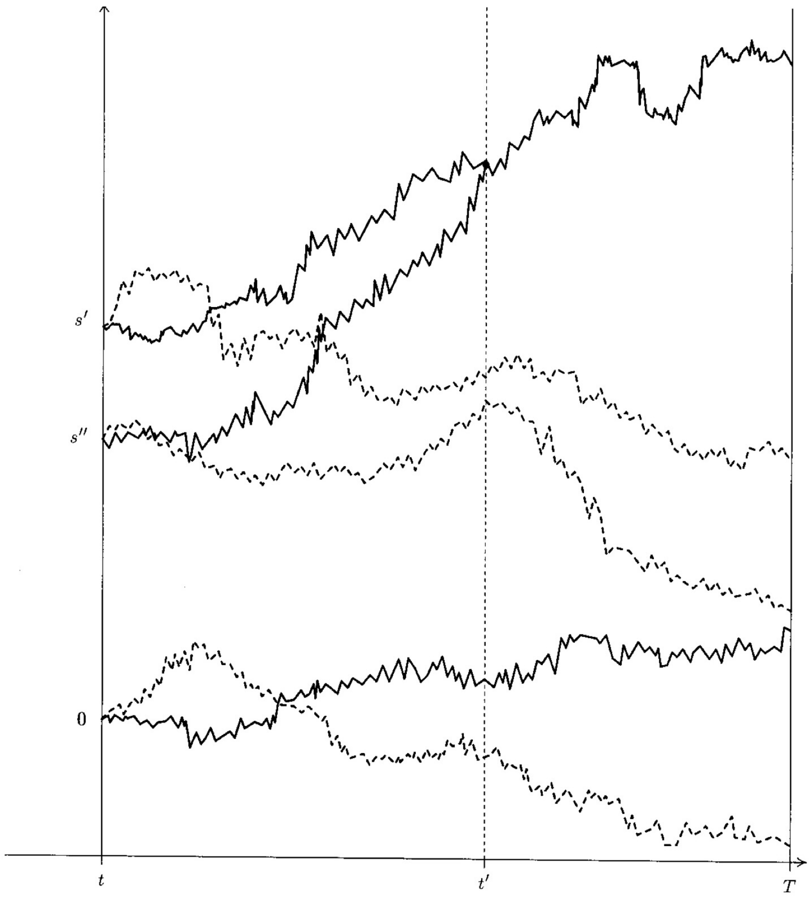
By (26), we must have then

$$x_1 < x_2 \implies X(t, x_1, \omega) < X(t, x_2, \omega) \quad \forall t \geqslant 0.$$

□

The illustration below, which can be found in [7, Figure 1], shows the non-crossing property in play described by Lemma (4) (both in the strong and weak sense).

Fig. 1: On the bottom, we observe two independent realizations of Brownian Motion (given by dense black and light black) up to time T. On the top, for each realization, we compare the paths starting at different positions at time t. For the dense black paths, we observe the weak formulation of the Non-Crossing Lemma (4): the paths eventually touch at time t' and stay the same afterwards. The strong formulation of the aforementioned lemma, by contrast, is demonstrated by the light dark paths: one always stays above the other.



2.5 Weak Solutions

Now, we will give a brief exposition into Weak solutions, since we will use them implicitly in due course in the following sections. Therefore, we present the following definition

Definition 11. *[1, Chapter 5, Section 3, Definition 16.1] The SDE:*

$$X_t = X_0 + \int_0^t \sigma(s, X_.)dB_s + \int_0^t b(s, X_.)ds \tag{27}$$

has a weak solution if for some $(\Omega, \mathcal{F}, \mathbb{P})$ satisfying the usual conditions and a given initial law μ, if there exist continuous semi-martingales X and B such that

(i)

$$X_t = X_0 + \int_0^t \sigma(s, X_.)\, dB_s + \int_0^t b(s, X_.)\, ds \tag{28}$$

(ii)

$$X_0 \textit{ has law } \mu; \tag{29}$$

(iii) For all $t \geqslant 0$,

$$\int_0^t \left(|\sigma(s, X_.)|^2 + |b(s, X_.)|\right) ds < \infty \tag{30}$$

(iv)

$$B = (B_t)_{t \geqslant 0} \textit{ is a Brownian motion adapted to } (\mathcal{F}_t) \tag{31}$$

Moreover, if there is a weak solution for every given probability measure μ*, then we say that the SDE* (27) *has a weak solution.*

This definition is only natural, insofar as it enables us to guarantee a solution of SDE (27) for, at least, one initial probability measure. It is clear, therefore, why weak solutions focus on the distributional properties of X. We may, for the purposes of simplicity, for example, fix an initial measure and study the problem in hand with full guarantee that a solution exists. It arises as well from the concept of a weak solution, that of a different notion of uniqueness. In the Strong setting, we were interested in path uniqueness, since a strong solution is by (4) driven only by the initial measurable process ξ and the adapted Brownian motion. In the Weak setting, however, we have the following appropriate notion of uniqueness

Definition 12. *[1, Chapter 5, Section 3, Definition 16.3][Uniqueness in Law] SDE* (27) *has uniqueness in law if for any two weak solutions* X *and* $\tilde{X}$ *with*

$$\mathcal{L}(X_0) = \mathcal{L}(\tilde{X}_0)$$

it holds that

$$\mathcal{L}(X) = \mathcal{L}(\tilde{X})$$

With this definition in hand, it is not so surprising that if we have pathwise uniqueness (4), then it immediately follows uniqueness of law. Indeed, taking into account Theorem (4), if we fix ξ to be a random variable governed by some probability measure μ, uniqueness of law easily follows by noticing that the strong solution is controlled by a Brownian motion and a deterministic function that applies to all set-ups. This is one of the consequences of the Yamada and Watanabe Theorem (9) stated below. But, before we state it, we will show a well-known example where not only both concepts of uniqueness are easily distinguished, but also completely justified.

Example 4. *[1, Chapter 5, Section 3, Example 16.5]* Consider a process $X_. : \Omega \to C([0,\infty) \times \mathbb{R})$ and respective one dimensional SDE

$$dX_t = f(X_t) dB_t,$$

where $f(x) = \mathbf{1}_{x \geqslant 0} - \mathbf{1}_{x<0}$. Suppose we have an initial setup where X_0 is governed by some law μ such that

$$X_t = X_0 + \int_0^t f(X_s) dB_s,$$

Then, notice that

$$[X, X]_t = \int_0^t f(X_s)^2 ds = t$$

which implies, by the Lévy´s characterization Theorem, that X_t follows the law of a Brownian motion. Then, under Definition (12), it follows uniqueness in law. Now, to prove existence, say that $X_t := \xi + B_t$. Consider

$$\tilde{B}_t = \int_0^t f(X_s) dB_s$$

again, by the Lévy´s characterization Theorem, $(\tilde{B}_t)_{t \geqslant 0}$ is a $\mathcal{F}_t$ Brownian Motion and observe that

$$\int_0^t f(X_s) d\tilde{B}_s = \int_0^t f(X_s)^2 dB_s = B_t$$

and thus $X_t = \xi + \int_0^t f(X_s) d\tilde{B}_s$ is a weak solution by Definition (12). But now notice that if

$$X_t = \int_0^t f(X_s) dB_s$$

that is, with $(\xi = 0)$, then we have

$$\int_0^t \mathbf{1}_{X_s=0} ds = 0$$

which, by Itô´s isometry, implies

$$\int_0^t \mathbf{1}_{X_s=0} dB_s = 0$$

It results that if X is a solution, then $-X$ is a solution as well. Although both have the same law, pathwise uniqueness as in (4), however, fails.

We are now prepared to state the Yamada and Watanabe Theorem, although with no proof. For that, we refer to [14].

Theorem 9 (Yamada and Watanabe Theorem)**.** *Let σ and b be predictable path functionals, and consider the SDE*

$$dX_t = \sigma(t, X.)\, dB_t + b(t, X.)\, dt. \tag{32}$$

Then SDE (32) *is exact as in* (5) *if and only if it has a weak solution and pathwise uniqueness holds.*

3 Robust Hedging under Uncertainty

3.1 Definitions and Initial Considerations

Fix $T > 0$ and assume, within this time horizon, that we trade continuously two assets in a market with no arbitrage. In particular, we assume that there exist an equivalent martingale measure (risk-neutral measure). Denote M_t the price of the money market at time t and a stock S_t the price of a stock at time t. More rigorously, we have the following hypothesis

Hypothesis 1. *[6, Section 2, Hypothesis 2.1] Assume no-arbitrage holds and let $\mathbb{Q}$ be a risk-neutral measure. Assume we have a setup $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t)_{t \geqslant 0}, B)$, which satisfies the usual conditions. Further, let $(M_t)_{t \geqslant 0}$ be a $\mathcal{F}_t$-adapted stochastic process associated with a deterministic function r_t (interest rate) such that*

$$M_t = \exp\Big(\int_0^t r_s ds\Big), \quad 0 \leqslant t \leqslant T \tag{33}$$

and a $(S_t)_{t \geqslant 0}$ a $\mathcal{F}_t$-adapted stochastic process associated with a predictable non-negative functional σ_t, the volatility process, governed by

$$dS_t = S_t \sigma_t dB_t + S_t r_t dt \tag{34}$$

where $B = (B_t)_{t \geqslant 0}$ is a one dimensional-Brownian motion adapted to $\mathcal{F}_t$. For simplicity, we may assume that $\mathcal{F} = \bigcup_{t=0}^{T} \mathcal{F}_t$ holds. Following the first Fundamental Theorem of Asset Pricing, we further have that S_t/M_t is a local Martingale with respect to $\mathbb{Q}$. In particular, we have

$$\frac{S_t}{M_t} = S_0 \exp\left[\int_0^t \sigma_u dB_u - \frac{1}{2}\int_0^t \sigma_u^2 du\right], \quad 0 \leqslant t \leqslant T \tag{35}$$

and we assume, in fact, that it is a square-integrable martingale, that is,

$$\mathbb{E}\left[\left(\frac{S_t}{M_t}\right)^2\right] < \infty, \quad 0 \leqslant t \leqslant T$$

Remark 7. Note that in Hypothesis (1), we do not assume that the volatility σ_t is not path-dependent. In fact, σ_t is assumed to be a very general predictable functional, which means that it may or may not depend only on its current price.

Remark 8. That $\frac{S_t}{M_t}$ is a local Martingale follows easily by noting that by Itô we have

$$d\left(\frac{S_t}{M_t}\right) = \frac{1}{M_t} dS_t - \frac{S_t r_t}{M_t} dt \tag{36}$$

$$= \frac{S_t \sigma_t}{M_t} dB_t + \frac{S_t r_t}{M_t} dt - \frac{S_t r_t}{M_t} dt = \frac{S_t \sigma_t}{M_t} dB_t \tag{37}$$

To show Equation (35), assume that $\frac{S_t}{M_t} = A_t$. Thus, (36) becomes

$$dA_t = A_t \sigma_t dB_t$$

with $A_0 = S_0$. By Itô, again, we then have

$$\begin{aligned} d(\log(A_t)) &= \frac{1}{A_t}dA_t \; - \frac{1}{2}\frac{1}{A_t^2}d[A,A]_t \\ &= \sigma_t dB_t - \frac{1}{2}\sigma_t^2 dt \end{aligned}$$

By integrating from 0 to t we obtain,

$$\log(A_t) = \log(A_0) + \int_0^t \sigma_u dB_u - \int_0^t \frac{1}{2}\sigma_u^2 du$$

and (35) follows. In addition, we assume that the market is not complete. In particular, that the filtration $\mathcal{F}$ is strictly larger than the one generated by the Brownian motion $(B_t)_{0\leqslant t\leqslant T}$. This choice will soon be clear, since the additional information in this case may account for volatility misspecification that cannot be replicated by trading in the underlying assets.

Due to the non-completeness of the market, we are forced to define a portfolio in the following way:

Definition 13. *[6, Section 2, Definition 2.3] Let $\{\theta_t : 0 \leqslant t \leqslant T\}$ be a bounded adapted process. Then, given an initial value $\Pi_\theta(0)$, the self financing process θ_t is governed by*

$$d\Pi_\theta = r(t)(\Pi_\theta(t) - S_t\theta_t)dt + \theta_t dS_t \tag{38}$$

where, by the same technique employed in Remark (8) applied to $\frac{\Pi_\theta}{M_t}$, we obtain

$$\Pi_\theta(t) = M(t)\Big[\Pi_\theta(0) + \int_0^t \theta_u d\left(\frac{S_u}{M_u}\right)\Big] \tag{39}$$

Remark 9. *Since $\frac{S_t}{M_t}$ is assumed to be a square-integrable martingale and θ_t is a bounded process by definition, it follows that $\frac{\Pi_{\theta(t)}}{M_t}$ is a square-integrable martingale.*

Definition 14. *[6, Section 2, Definition 2.5] A payoff function is a convex function h, defined on $(0,\infty)$, and having bounded one-sided derivatives, that is,*

$$|h'(x\pm)| \leqslant C \quad \forall x > 0. \tag{40}$$

Definition 15. *[15, Part III, Section 7.3, Definition 7.7] A vanilla contingent claim is a contract that pays $h(S_T)$ at time T.*

Example 5. *Let $h(x) = (x - K)^+$, for a fixed price K.*

European call option:

$$C_{call} = (S_T - K)^+$$

where the owner of a European call option has the right but not the obligation to buy the underlying asset with price S_T at time T for a fixed strike price K.

European put option:

$$C_{put} = (K - S_T)^+$$

where the owner of a European put option has the right but not the obligation to sell the underlying asset with price S_T at time T for a fixed strike price K.

Path-dependent contingent claim (e.g., up-and-out call): Let $B > 0$ be a fixed barrier level. Then,

$$C_{u\&o}^{call} = \begin{cases} (S_T - K)^+, & \text{if } \max_{0 \leqslant t \leqslant T} S_t < B, \\ 0, & \text{otherwise.} \end{cases}$$

Definition 16. *[15, Part III, Subsection 7.3, Assumption 7.3.1] An adapted stochastic process $\{V(t, S_t) : 0 \leqslant t \leqslant T\}$ is called the price process for a vanilla contingent claim if*

$$V(T, S_T) = h(S_T), \text{ a.s.}, \tag{41}$$

and $(S_t, V(t, S_t))$ admits no-arbitrage.

In a non-complete market, we do not have an explicit expression for the price process, but in a complete one, by virtue of the second fundamental Theorem of asset pricing, under the unique risk-neutral measure $\mathbb{Q}$, it follows immediately by Definition (16) and the martingale property that

$$\frac{V(t, S_t)}{M_t} = \mathbb{E}_{\mathbb{Q}}\left[\frac{V(T, S_T)}{M_T} \,\middle|\, \mathcal{F}_t\right] = \mathbb{E}_{\mathbb{Q}}\left[\frac{h(S_T)}{M_T} \,\middle|\, \mathcal{F}_t\right] \tag{42}$$

and so, we are able to replicate directly the option using our self-financing portfolio. We will provide more details on this soon. This leads us to the following definition.

Definition 17. *[6, Section 2, Definition 2.6] Suppose you have some self-financing portfolio Π_θ, governed by the strategy $\{\theta_t : 0 \leqslant t \leqslant T\}$ defined on the finite time horizon T. Define the error process $e_t := \Pi_\theta(t) - V(t, S_t)$, where $V(t, S_t)$ represents the price process of some contingent claim. If we assume further that $\Pi_\theta(0) = V(0, S_0)$ then the pair $(\theta_t, V(t, S_t))_{0 \leqslant t \leqslant T}$ is said to be*

1. *a replicating strategy, if $\frac{e_t}{M_t}$ is identically equal to 0,*
2. *a super-strategy, if $\frac{e_t}{M_t}$ is non-decreasing,*
3. *a sub-strategy, if $\frac{e_t}{M_t}$ is non-increasing.*

Remark 10. If the pair $(\theta_t, V(t, S_t))$ is a super-strategy, we are guaranteed by definition to have $e_T \geqslant 0$. Since $\frac{\Pi_\theta(t)}{M_t}$ is a martingale by (39) and $\frac{e_t}{M_t}$ is non-decreasing by assumption, it follows that $\frac{V(t, S_T)}{M_t}$ is a super martingale and we have the following relation

$$\frac{V(t, S_t)}{M_t} \geqslant \mathbb{E}\left[\frac{h(S_T)}{M_T} \,\middle|\, \mathcal{F}_t\right]$$

Thus, in particular, given that $M_0 = 1$, we have

$$V(0, S_0) \geqslant \mathbb{E}\left[\frac{h(S_T)}{M_T}\right]$$

Under the assumptions of no-arbitrage and market completeness, the quantity

$$\mathbb{E}\left[\frac{h(S_T)}{M_T}\right]$$

represents the unique arbitrage-free price of the contingent claim. Indeed, a replicating strategy attains the payoff $h(S_T)$ exactly with this initial capital, while any super-hedging strategy must require at least as much initial wealth. Otherwise, if an investor could super-replicate the claim with strictly smaller initial capital, then purchasing the claim below its replicating price and hedging it accordingly would generate a risk-free profit, thereby creating an arbitrage opportunity.

3.2 Classical Black-Scholes

Now, let us focus more precisely on the price process under model (34). Suppose we have a function $f(t, S_t) := V(t, S_t)$ and, for simplicity, further assume that S_t is already discounted and that $\sigma_t = \sigma$, i.e, a constant. By Itô, we have

$$\begin{aligned} df(t, S_t) &= f_t(t, S_t)dt + f_s(t, S_t)dS_t + f_{ss}(t, S_t)d[S, S]_t \\ &= (f_t(t, S_t) + f_{ss}(t, S_t)\sigma_t^2 S_t^2)dt + f_s(t, S_t)dS_t, \quad 0 \leqslant t \leqslant T \end{aligned}$$

where $[.\,,.]$ denotes the quadratic variation between 2 processes. By definition, the price process $V(t, S_t)$ is a martingale. Therefore, we must have

$$f_t(t, S_t) + f_{ss}(t, S_t)\sigma_t^2 S_t^2 = 0, \quad 0 \leqslant t \leqslant T$$

with the extra condition that

$$V(T, S_T) = h(S_T)$$

Further, since $V(t, S_t)$ is a martingale, it follows that

$$dV(t, S_t) = f_s(t, S_t)dS_t, \quad 0 \leqslant t \leqslant T$$

or in the integral form:

$$V(t, S_t) = V(0, S_0) + \int_0^t f_s(u, S_u)dS_u, \quad 0 \leqslant t \leqslant T$$

That is, with a fixed initial price $V(0, S_0)$, we are able to replicate the option $h(S_T)$ at expiration if and only if we know explicitly $f_s(t, S_t) = \frac{\partial}{\partial s}V(t, S_t)$. Let us look more closely into it. First, notice that, under the taken assumptions, we have

$$\begin{aligned} S_T &= S_0 \exp(\sigma B_T - \frac{1}{2}\sigma^2 T) = S_0 \exp(\sigma B_t - \frac{1}{2}\sigma^2 t) \exp(\sigma(B_T - B_t) - \frac{1}{2}\sigma^2(T - t)) \\ &= S_t \exp(\sigma(B_T - B_t) - \frac{1}{2}\sigma^2(T - t)) \end{aligned}$$

Thus, we have, by definition

$$V(t, S_t) = \mathbb{E}[h(S_T) \,|\, \mathcal{F}_t] = \mathbb{E}[h(S_t \exp(\sigma(B_T - B_t) - \frac{1}{2}\sigma^2(T - t))) \,|\, \mathcal{F}_t] \tag{43}$$

Now, notice that since the filtration $\mathcal{F}_t$ is generated by B_t, by the independent increments property of the Brownian Motion, $B_T - B_t$ is, therefore, independent of $\mathcal{F}_t$. As a result, (43) becomes

$$\int_{-\infty}^{\infty} h(s \exp(\sigma y - \frac{1}{2}\sigma^2(T - t)))d\gamma_{T-t}$$

where $\gamma_{T-t} \stackrel{d}{=} \mathcal{N}(0, T - t)$, i.e, a normal distribution with mean 0 and variance $T - t$. Further, it becomes

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} h(s \exp(\sigma y(\sqrt{T - t}) - \frac{1}{2}\sigma^2(T - t)))e^{-\frac{y^2}{2}} dy \tag{44}$$

For the sake of simplicity, let us assume that we are dealing with an European call option with strike price K, that is, $h(S_T) = (S_T - K)^+$. Then, set $\alpha = \sigma\sqrt{T-t}$ and the integral (44) becomes

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} (se^{(y\alpha - \frac{1}{2}\alpha^2)} - K)^+ e^{-\frac{y^2}{2}} dy \tag{45}$$

Accordingly, set

$$u = \frac{-\log\left(\frac{s}{K}\right) + \frac{1}{2}\alpha^2}{\alpha} \tag{46}$$

then, we have

$$se^{(\sigma y\alpha - \frac{1}{2}\alpha^2)} \geqslant K \implies y \geqslant u$$

And (45) results in

$$\begin{aligned}
&= \frac{1}{\sqrt{2\pi}} \int_{u}^{\infty} (se^{(\sigma y\alpha - \frac{1}{2}\alpha^2)} - K) e^{-\frac{y^2}{2}} dy \\
&= \frac{s}{\sqrt{2\pi}} \int_{u}^{\infty} e^{(\sigma y\alpha - \frac{1}{2}\alpha^2 - -\frac{y^2}{2})} dy - K\Phi(-u) \\
&= \frac{s}{\sqrt{2\pi}} \int_{u}^{\infty} e^{-\frac{(y-\alpha)^2}{2}} dy - K\Phi(-u) \\
&= s\Phi(\alpha - u) - K\Phi(-u)
\end{aligned} \tag{47}$$

where Φ is the probability distribution of a Normal Distribution with mean 0 and variance 1.

Thus, the replicating strategy $\Delta(t, S_t)$ is given by

$$\Delta(t, S_t) = \frac{\partial}{\partial s} V(t, S_t) = \Phi(\alpha - u) + s\frac{\partial}{\partial s}\Phi(\alpha - u) - K\frac{\partial}{\partial s}\Phi(-u), \tag{48}$$

where u is defined in (46).

By construction, $\alpha - u = -u + \sigma\sqrt{T-t}$, hence

$$\frac{\partial}{\partial s}(\alpha - u) = \frac{\partial}{\partial s}(-u).$$

Moreover,

$$(\alpha - u)^2 = u^2 - 2\alpha u + \alpha^2 = u^2 + 2\log\left(\frac{s}{K}\right),$$

which implies

$$-\frac{1}{2}(\alpha - u)^2 + \frac{1}{2}u^2 = \log\left(\frac{K}{s}\right).$$

Therefore,

$$\frac{\frac{\partial}{\partial s}\Phi(\alpha - u)}{\frac{\partial}{\partial s}\Phi(-u)} = \exp\left(-\frac{1}{2}(\alpha - u)^2 + \frac{1}{2}u^2\right) = \frac{K}{s}.$$

Substituting into (48) yields

$$\Delta(t, S_t) = \Phi(\alpha - u).$$

Finally, defining

$$d^+ = \frac{\log\left(\frac{s}{K}\right) + \frac{1}{2}\alpha^2}{\alpha},$$

we obtain

$$\Delta(t, S_t) = \Phi(d^+). \tag{49}$$

Thus, in this setting, we obtain an explicit and relatively easy representation of the price process defined in (16), together with the associated self-financing replicating strategy given by (49). Moreover, this example illustrates the completeness of the Black-Scholes model, in the sense that any European contingent claim can be replicated by trading in the underlying asset.

3.3 Convexity of European Contingent Claims

Now, let us assume that the tradable asset S_t is governed by SDE (34), while we believe instead that it is governed by some volatility $\bar{\sigma}(t, S_t)$, that is,

$$dS_t = S_t\bar{\sigma}(t, S_t)dB_t + S_t r dt \tag{50}$$

with r as in (33). Notice that, by Example provided in [6, 4. Stochastic Volatility Counterexample], we assume that the misspecified volatility is stochastic only through S_t. First of all, we state the following Hypothesis.

Hypothesis 2. *[6, Section 5, Hypothesis 5.1] Suppose $\bar{\sigma} : [0, T] \times (0, \infty) \to \mathbb{R}$ is a continuous bounded function, for which $\frac{\partial}{\partial s}(s\bar{\sigma}(t, s))$ is Lipschitz continuous with respect to s and uniformly continuous with respect to t.*

The following lemma will prove exceedingly useful for the main theorem of this subsection, because we obtain, together with Hypothesis (1), the so desired property of convexity for the price process defined weakly in (42). We will prove this result under the further assumption that the price process starts at time 0, although it is possible to prove it for other times within the finite horizon $(0, T)$. For that particularly, I refer to [7]. Therefore, following [6], we will consider, under misspecified model (50), the price process $v_{\bar{\sigma}}$ starting at time 0,

$$v_{\bar{\sigma}} := \frac{1}{M_T}\mathbb{E}[h(S_T)] \tag{51}$$

Lemma 6. *[6, Section 5, Theorem 5.2] Under Hypothesis* (2) *and the ruling dynamics described in* (50)*, the contingent European claim $v_{\bar{\sigma}}$ is convex with respect to S_0.*

Proof. Consider the SDE (50). By Theorem (8), we can find a modification of S_t, called again S_t, such that for each $t \in [0, T]$ and for each realization path w, the map $x \to S_t(w)$ is a C^1-diffeormophism. Set S_t^x to be this modification with the additional property that $S_0^x = x$. It follows that the process $D_t^x = \frac{\partial}{\partial x}(S_t^x)$, by the Chain Rule, is given by

$$\begin{aligned}
dD_t^x &= D_t^x r_t dt + \frac{\partial}{\partial x}(S_t^x\bar{\sigma}(t, S_t))dB_t \\
&= D_t^x r_t dt + D_t^x\frac{\partial}{\partial s}(\rho(t, S_t))dB_t \\
&= D_t^x(r_t dt + \frac{\partial}{\partial s}(\rho(t, S_t))dB_t)
\end{aligned}$$

where $\rho(t,s) = s\bar{\sigma}(t,s)$ is well-defined. Moreover, we immediately obtain that

$$\frac{D_t^x}{M_t} = \zeta_t^x,$$

where

$$\zeta_t^x = \exp\left(\int_0^t \frac{\partial}{\partial s}\rho(u,S_u^x)\,dB_u - \frac{1}{2}\int_0^t \left(\frac{\partial}{\partial s}\rho(u,S_u^x)\right)^2 du\right),$$

is a strictly positive martingale.

Define a new probability measure $\mathbb{Q}^x$ on $(\Omega,\mathcal{F})$ by

$$\frac{d\mathbb{Q}^x}{d\mathbb{Q}} = \zeta^x(T).$$

According to Girsanov's Theorem, under $\mathbb{Q}^x$ the process

$$B_t^x = B_t - \int_0^t \frac{\partial}{\partial s}\rho(u,S_u^x)\,du \tag{52}$$

is a Brownian motion. Now, since

$$\frac{S_t^x}{M_t} = x + \int_0^t \frac{\rho(u,S_u^x)}{M_u}dB_u$$

for $y > x$, we have

$$\frac{S_t^y}{M_t} - \frac{S_t^x}{M_t} = y - x + \int_0^t \frac{\rho(u,S_u^y) - \rho(u,S_u^x)}{M_u}dB_u$$

for all $0 \leqslant t \leqslant T$. Thus, it follows

$$\mathbb{E}\left[\left(\frac{S_t^y - S_t^x}{M_t}\right)^2\right] \leqslant K\left[(y-x)^2 + \mathbb{E}\left[\int_0^t \left(\frac{\rho(u,S_u^y) - \rho(u,S_u^x)}{M_u}\right)dB_u\right]^2\right] \tag{53}$$

for a constant K. The cross term vanishes under expectation, since the integral is a Martingale. By Itô´s Isometry and the Lipschitz condition with respect to s assumed in Hypothesis (2), for some constant C, (53) becomes

$$\mathbb{E}\left[\left(\frac{S_t^y - S_t^x}{M_t}\right)^2\right] \leqslant C\left[(y-x)^2 + \mathbb{E}\left[\int_0^t \left(\frac{S_u^y - S_u^x}{M_u}\right)^2 du\right]\right]$$

for all $t \in (0,T]$. Thus, as a consequence, by Grönwall´s Lemma (2), it results in

$$\mathbb{E}\left[(S_t^x - S_t^y)^2\right] \leqslant CM_t^2(y-x)^2 e^{KT}, \quad 0 \leqslant t \leqslant T$$

or, more plainly, for a global constant, say C again

$$\mathbb{E}\left[(S_t^x - S_t^y)^2\right] \leqslant C(y-x)^2, \quad 0 \leqslant t \leqslant T \tag{54}$$

Therefore, for $y > x > 0$, consider the following,

$$\frac{v_{\bar{\sigma}}(y) - v_{\bar{\sigma}}(x)}{y - x} = \frac{1}{M_T(y-x)} \mathbb{E}\left[h(S_T^x) - h(S_T^y)\right]$$

Since $y > x$, we know by the Non - Crossing Lemma (4) that $S_t^y > S_t^x$ almost surely (although a strict inequality is too strong). Thus, since (54) and (14) hold, the difference quotient is uniformly integrable in L^2 and, by the dominated convergence theorem, we have

$$\begin{aligned}
&\limsup_{y \to x^+} \frac{v_{\bar{\sigma}}(y) - v_{\bar{\sigma}}(x)}{y - x} \leqslant \mathbb{E}\left[h'(S_T^{x+}) \limsup_{y \to x^+} \frac{(S_t^x - S_t^y)}{M_T(y-x)}\right] \\
&= \mathbb{E}[h'(S_T^{x+})\zeta_t^x] = \mathbb{E}^x[h'(S_T^{x+})]
\end{aligned}$$

where $\mathbb{E}^x[.]$ denotes the expectation under $\mathbb{Q}^x$. Similarly, we have

$$\liminf_{y \to x^+} \frac{v_{\bar{\sigma}}(y) - v_{\bar{\sigma}}(x)}{y - x} \leqslant \mathbb{E}^x[h'(S_T^{x+})]$$

The same reasoning applies to $\mathbb{E}^x[h'(S_T^{x-})]$. As a result, we have

$$v'_{\bar{\sigma}}(x) = \mathbb{E}^x[h'(S_T^x)]$$

Now, notice that the SDE (50) can be rewritten in the form

$$dS_t^x = (S_t^x r_t + \rho(t, S_t))dt + \rho(t, S_t)dB_t^x$$

where B_t^x denotes the transformed Brownian Motion given by (52). If we consider a weak solution of this SDE, by the boundness of $\bar{\sigma}$ and Lipschitz condition with respect to s, we know uniqueness in law (12) holds. Therefore, we can take another process $\tilde{S}_t^x$ of the following form

$$d\tilde{S}_t^x = (\tilde{S}_t^x r_t + \rho(t, \tilde{S}_t))dt + \rho(t, \tilde{S}_t)dB_t$$

where, B_t denotes the original Brownian Motion under $\mathbb{P}$ and such that $\tilde{S}_0^x = x$. Using uniqueness in law, we obtain

$$v'_{\bar{\sigma}}(x) = \mathbb{E}[h'(\tilde{S}_T^x)],$$

and using again the No-Crossing Lemma (4), we obtain the increasingness of this function. □

Corollary 5. *Under Hypothesis* (2) *and the convexity of* h, *the European contingent claim satisfies*

$$|v_{\bar{\sigma}}(x)| \leqslant C, \quad \textit{for all } x > 0$$

Proof. It follows by Lemma (6) and the convexity of h. □

3.4 Volatility Misspecification and Main Theorem

We are thus ready to state the main theorem of this section, which will be important for future considerations.

For that, we first assume that the asset S_t is wrongfully assumed by the investor to be governed by a misspecified volatility $\bar{\sigma}_t$, as in (50) while it is actually governed by the SDE (34), with true volatility σ. Naturally, if the misspecified volatility happens to be equal to the true one, that is $\bar{\sigma} = \sigma$, almost surely, the market would be complete and, as a consequence, one would be in the case of a replicating strategy, that is, one could replicate the option risk-free. But in a practical setting, through local volatility models, the estimate $\bar{\sigma}$ rarely represents the true volatility σ, because we are only assuming one source of randomness.

Theorem 10. *[6, Section 6, Theorem 6.2] Suppose that Hypothesis (2) holds and assume that we believe the price process is governed by (50), while the true model governing the asset is given by (34). Further, assume that both the deterministic interest rate function $r : [0,T] \to \mathbb{R}$ and $\bar{\sigma}_t : [0,T] \times (0,\infty) \to \infty$ are Hölder Continuous, as in (10). Consider the self-financing portfolio process Π_θ governed by (38), with initial investment*

$$\Pi_\theta(0) = V(0, S_0),$$

and hedging strategy

$$\theta_t = \frac{\partial}{\partial s} V(t, S_t).$$

Then, if we have

$$\sigma_t \leqslant \bar{\sigma}_t(t, S_t), \text{ almost surely}$$

for all $t \in [0,T]$, the pair (θ_t, V_t) is a superstrategy and we have

$$\Pi_\theta(T) \geqslant h(S(T)) \tag{55}$$

and

$$V(0, S_0) \geqslant \mathbb{E}\left[\frac{h(S_T)}{M_T}\right] \tag{56}$$

On the other hand, if

$$\sigma_t \geqslant \bar{\sigma}_t(t, S_t), \text{ almost surely}$$

holds for all $t \in [0,T]$, then the pair $(\theta_t, V(t, S_t))$ is a substrategy and inequalities (55) and (56) are reversed.

For this, we prove a Lemma and a Proposition that hold without any strict assumption.

Lemma 7. *Consider a European contingent claim priced incorrectly according to model (50). More precisely, set correspondingly $v_{\bar{\sigma}}(t,x)$ to be the initial price of the incorrect European contingent claim - starting, more generally, at time t associated with stock S_t starting at price x - according to dynamics (50) as*

$$v_{\bar{\sigma}}(t,x) = \mathbb{E}[e^{-(\int_t^T r_s ds)} h(S_T^{t,x})], \quad 0 \leqslant t \leqslant T, \ x > 0. \tag{57}$$

so that starting at time $t = 0$, it coincides with definition (51). Therefore, if we would have $\bar{\sigma} = \sigma$ almost surely, we would be on the case of a replicating strategy as in (17). Then, we have the following

$$\mathcal{L}(v_{\bar{\sigma}}(t,x)) = 0, \tag{58}$$

where the operator $\mathcal{L}$ is defined for $f \in C^{1,2}([0,T) \times (0,\infty))$ as

$$\mathcal{L}(f(t,x)) = r_t f(t,x) - \frac{\partial}{\partial t} f(t,x) - r_t x \frac{\partial}{\partial x} f(t,x) - \frac{1}{2}\bar{\sigma}_t(x)^2 x^2 \frac{\partial^2}{\partial x^2} f(t,x)$$

for all $0 \leqslant t < T$.

Proof. To ease notation, fix (t, x) and define

$$X_u = S_u^{t,x}, \qquad u \geqslant t.$$

as a stochastic process with the same dynamics as S_t in (50), starting at time t with associated price S_t^x. Thus, we are generalizing the starting time $(t = 0)$, to the whole time horizon $(0, T]$. Define, to that end, the discount factor

$$\tilde{M}_u = \exp\left(-\int_t^u r_s ds\right).$$

Now, consider the process

$$Y_u = \tilde{M}_u v_{\bar{\sigma}}(u, X_u).$$

Again, we assume that the underlying asset is wrongfully governed by $\bar{\sigma}$. Thus, by assuming further completeness of the market, we impose the martingale condition. In our present notation, this means that

$$Y_u = \tilde{M}_u v_{\bar{\sigma}}(u, X_u) = \mathbb{E}[\tilde{M}_T h(S_T) \mid \mathcal{F}_t] \tag{59}$$

is a Martingale. Since $v_{\bar{\sigma}} \in C^{1,2}$, Itô's formula implies

$$dv_{\bar{\sigma}}(u, X_u) = \frac{\partial v_{\bar{\sigma}}}{\partial t}(u, X_u)du + \frac{\partial v_{\bar{\sigma}}}{\partial x}(u, X_u)dX_u + \frac{1}{2}\frac{\partial^2 v_{\bar{\sigma}}}{\partial x^2}(u, X_u)d[X_u, X_u]$$

Since, by (50), we have

$$dX_u = r_u X_u du + \bar{\sigma}(u, X_u)X_u dB_u,$$

and

$$d[X, X]_u = \bar{\sigma}(u, X_u)^2 X_u^2 du,$$

we obtain,

$$\begin{aligned} dv_{\bar{\sigma}}(u, X_u) = &\Big[\frac{\partial v_{\bar{\sigma}}}{\partial t} + r_u X_u \frac{\partial v_{\bar{\sigma}}}{\partial x} + \frac{1}{2}\bar{\sigma}(u, X_u)^2 X_u^2 \frac{\partial^2 v_{\bar{\sigma}}}{\partial x^2}\Big]du \\ &+ \bar{\sigma}(u, X_u)X_u \frac{\partial v_{\bar{\sigma}}}{\partial x} dB_u. \end{aligned}$$

Since

$$d\tilde{M}_u = -r_u \tilde{M}_u du,$$

the product rule yields

$$\begin{aligned} dY_u &= d(\tilde{M}_u v_{\bar{\sigma}}(u, X_u)) \\ &= \tilde{M}_u \, dv_{\bar{\sigma}}(u, X_u) - r_u \tilde{M}_u v_{\bar{\sigma}}(u, X_u)du. \end{aligned}$$

Substituting the Itô expansion,

$$\begin{aligned} dY_u = \tilde{M}_u &\Big[\frac{\partial v_{\bar{\sigma}}}{\partial t} + r_u X_u \frac{\partial v_{\bar{\sigma}}}{\partial x} + \frac{1}{2}\bar{\sigma}(u, X_u)^2 X_u^2 \frac{\partial^2 v_{\bar{\sigma}}}{\partial x^2} - r_u v_{\bar{\sigma}}\Big]du \\ &+ \tilde{M}_u \bar{\sigma}(u, X_u)X_u \frac{\partial v_{\bar{\sigma}}}{\partial x}(u, X_u)dB_u. \end{aligned}$$

By the observation (59) $\{Y_u\}_{u \geqslant t}$ is a martingale, thus the drift coefficient must vanish, which implies

$$\frac{\partial v_{\bar{\sigma}}}{\partial t} + r_t x \frac{\partial v_{\bar{\sigma}}}{\partial x} + \frac{1}{2}\bar{\sigma}(t,x)^2 x^2 \frac{\partial^2 v_{\bar{\sigma}}}{\partial x^2} - r_t v_{\bar{\sigma}} = 0, \quad t < T.$$

Finally, at $t = T$,

$$v_{\bar{\sigma}}(T,x) = h(x).$$

The conclusion of (10) follows. $\square$

Within this framework, by incorrectly assuming dynamics (42), we now set $v_{\bar{\sigma}}(t, S_t)$ to be the price of the European contingent claim and let $\bar{\Delta} = \frac{\partial}{\partial x} v_{\bar{\sigma}(t,x)}$ be our hedging strategy. With initial price $v_{\bar{\sigma}}(0, S_0)$, the self-financing portfolio $\Pi_{\bar{\Delta}}$ will evolve according to the dynamics prescribed in (38). We have the following Proposition.

Proposition 5. *Under the true dynamics* (34) *of* S_t, *let* $\Pi_{\bar{\Delta}}$ *be the (incorrect) portfolio process described above and the corresponding error process* $e_{\bar{\Delta}}$ *defined as*

$$e_{\bar{\Delta}} = \Pi_{\bar{\Delta}}(t, S_t) - v_{\bar{\sigma}}(t, S_t)$$

where $v_{\bar{\sigma}}(t, S_t)$ *is given by* (57). *Then, it holds*

$$e_{\bar{\Delta}} = M_t \int_0^t \frac{1}{2M_u}[\bar{\sigma}(u, S_u)^2 - \sigma_u^2] S_u^2 \frac{\partial^2}{\partial x^2} v_{\bar{\sigma}}(u, S_u)\, du$$

Proof. First, notice that by Itô and by reminding ourselves that S_t is governed by the true volatility σ_t, that is, by SDE (34), we have

$$d(v_{\bar{\sigma}}(t, S_t)) = \frac{\partial}{\partial t} v_{\bar{\sigma}}(t, S_t) dt + \frac{\partial}{\partial x} v_{\bar{\sigma}}(t, S_t) dS_t + \frac{1}{2}\frac{\partial^2}{\partial x^2} v_{\bar{\sigma}}(t, S_t) S_t^2 \sigma_t^2 dt \tag{60}$$

Only by incorrectly pricing the contingent claim, under (50), through which the PDE (58) holds, do we obtain

$$\frac{\partial}{\partial t} v_{\bar{\sigma}}(t, S_t) = r_t v_{\bar{\sigma}}(t, S_t) - r_t S_t \frac{\partial}{\partial x} v_{\bar{\sigma}}(t, S_t) - \frac{1}{2}\bar{\sigma}(t, S_t)^2 S_t^2 \frac{\partial^2}{\partial x^2} v_{\bar{\sigma}}(t, S_t)$$

and thus, (60) becomes

$$\begin{aligned} d(v_{\bar{\sigma}}(t, S_t)) &= r_t v_{\bar{\sigma}}(t, S_t) dt - r_t S_t \bar{\Delta} dt - \frac{1}{2}\bar{\sigma}(t, S_t)^2 \frac{\partial^2}{\partial x^2} v_{\bar{\sigma}}(t, S_t) dt \\ &\quad + \bar{\Delta} dS_t + \frac{1}{2}\frac{\partial^2}{\partial x^2} v_{\bar{\sigma}}(t, S_t) S_t^2 \sigma_t^2 dt \\ &= r_t v_{\bar{\sigma}}(t, S_t) dt - r_t S_t \bar{\Delta} dt + \bar{\Delta} dS_t \\ &\quad + \frac{1}{2}\frac{\partial^2}{\partial x^2} v_{\bar{\sigma}}(t, S_t) S_t^2 (\sigma_t^2 - \bar{\sigma}(t, S_t)^2) dt \end{aligned} \tag{61}$$

Now, consider the discounted tracking error. By the product rule, we must have

$$d\left(\frac{e_{\bar{\Delta}}(t)}{M_t}\right) = d(\Pi_{\bar{\Delta}}(t, S_t) - v_{\bar{\sigma}}(t, S_t))) \frac{1}{M_t} + (\Pi_{\bar{\Delta}}(t, S_t) - v_{\bar{\sigma}}(t, S_t)) d\left(\frac{1}{M_t}\right)$$

which by (61) and (38) becomes

$$
\begin{aligned}
&\Big(r_t\Pi_{\bar{\Delta}}(t,S_t)dt + \bar{\Delta}dS_t - \bar{\Delta}r_tS_tdt - r_tv_{\bar{\sigma}}(t,S_t)dt - \bar{\Delta}dS_t \\
&+ \bar{\Delta}r_tS_tdt + \frac{1}{2}\frac{\partial^2}{\partial x^2}v_{\bar{\sigma}}(t,S_t)S_t^2(\bar{\sigma}(t,S_t)^2 - \sigma_t^2)dt\Big)\frac{1}{M_t} \\
&- \frac{r_t}{M_t}\Pi_{\bar{\Delta}}(t,S_t)dt + \frac{r_t}{M_t}v_{\bar{\sigma}}(t,S_t)dt \\
&= \frac{1}{2M_t}\frac{\partial^2}{\partial x^2}v_{\bar{\sigma}}(t,S_t)S_t^2(\bar{\sigma}(t,S_t)^2 - \sigma_t^2)dt
\end{aligned}
$$

This, together with $\Pi_{\bar{\Delta}}(0,S_0) = v_{\bar{\sigma}}(0,S_0))$ implies that

$$
\frac{e_{\bar{\Delta}}(t)}{M_t} = \int_0^t \frac{1}{2M_u}\frac{\partial^2}{\partial x^2}v_{\bar{\sigma}}(u,S_u)S_u^2(\bar{\sigma}(u,S_u)^2 - \sigma_u^2)\,du
$$

□

Proof of Theorem (10). Using the hedge $\bar{\Delta}$, by the convexity of $v_{\bar{\sigma}}(u,S_u)$ proved in (6) and Proposition (5), we obtain

$$
e_{\bar{\Delta}}(t) \geqslant 0
$$

Thus, $\bar{\Delta}$ is a superstrategy according to (17) and the subsequent conclusions follow by Remark (10). □

3.5 Example

We follow [6, Section 11, Arithmetic Mean Example 11.1]. Let σ_1 and σ_2 be constants. Define S_t and S'_t such that

$$
dS_t = \sigma_1 S_t dB_t, \qquad dS'_t = \sigma_2 S'_t dB'_t,
$$

where B_t and B'_t are independent Brownian motions.

Further, define

$$
S''_t = \frac{S_t + S'_t}{2}.
$$

which is governed by

$$
dS''_t = \frac{1}{2}\left[S_t\sigma_1 dB_t + S'_t\sigma_2 dB'_t\right]
$$

Accordingly, set

$$
B''_t = \int_0^t \frac{\sigma_1 S_u}{\sqrt{\sigma_1^2 S_u^2 + \sigma_2^2 S_u'^2}}dB_u + \int_0^t \frac{\sigma_2 S'_u}{\sqrt{\sigma_1^2 S_u^2 + \sigma_2^2 S_u'^2}}dB'_u,
$$

so that B''_t is a Brownian motion.

To verify this, using Itô's isometry Lemma and the independence of the Brownian Motions, notice that

$$
\begin{aligned}
d[B''_t, B''_t] &= \frac{\sigma_1^2 S_u^2}{\sigma_1^2 S_u^2 + \sigma_2^2 S_u'^2}dt + \frac{\sigma_2^2 S_u'^2}{\sigma_1^2 S_u^2 + \sigma_2^2 S_u'^2}dt \\
&+ \frac{\sigma_1 S_u}{\sigma_1^2 S_u^2 + \sigma_2^2 S_u'^2}\frac{\sigma_2 S'_u}{\sigma_1^2 S_u^2 + \sigma_2^2 S_u'^2}d[B_t, B'_t] \\
&\implies [B''_t, B''_t] = t
\end{aligned}
$$

By Lévy´s characterization Theorem, B''_t is then a Brownian Motion. Now, set accordingly

$$dS''_t = S''_t \sigma_t dB''_t$$

with σ_t, now stochastic, given by

$$\sigma_t = \frac{\sqrt{\sigma_1^2 S_t^2 + \sigma_2^2 S_t'^2}}{S_t + S'_t} \tag{62}$$

If we are interested in pricing the European call on S''_t with strike price K at time T, $\mathbb{E}[(S''_T - K)^+]$ is not easy to compute. Therefore, if we find constants $\sigma_{\min}$ and $\sigma_{\max}$ satisfying

$$\sigma_{\min} \leqslant \sigma_t \leqslant \sigma_{\max}$$

for all $t \in [0, T]$ almost surely, then, according to Theorem (10), it follows that the initial price associated with the option $v_{\sigma_t}(0, S''_0)$, by (47), satisfies

$$S''_0 \phi(\alpha_{\min} - u_{\min}) - K\phi(-u_{\min}) \leqslant v_{\sigma_t}(0, S''_0) \leqslant S''_0 \phi(\alpha_{\max} - u_{\max}) - K\phi(-u_{\max}) \tag{63}$$

where $\alpha_{\min} = \sigma_{\min}\sqrt{T}$, $\alpha_{\max} = \sigma_{\max}\sqrt{T}$ and accordingly

$$u_{\min} = \frac{-\log\left(\frac{S''_0}{K}\right)}{\alpha_{\min}} + \frac{1}{2}\alpha_{\min}, \quad u_{\max} = \frac{-\log\left(\frac{S''_0}{K}\right)}{\alpha_{\max}} + \frac{1}{2}\alpha_{\max}$$

where the precise calculations of these quantities are accounted in Subsection (3.2). One can set

$$\sigma_{\min} = \frac{\sigma_1\sigma_2}{\sqrt{\sigma_1^2 + \sigma_2^2}}, \quad \sigma_{\max} = \max(\sigma_1, \sigma_2)$$

Let us verify both inequalities. With regard to $\sigma_{\max}$, one first notices that σ_t is scale invariant. Therefore, set $x = S_t$ and $y = S'_t$. Then, it is enough to consider the subspace $\Omega \subset \mathbb{R}^+ \times \mathbb{R}^+$ for which

$$x + y = 1$$

holds. Then, we have

$$\sigma_t \leqslant \sqrt{\sigma_{\max}^2 (x^2 + (1-x)^2)} \leqslant \sigma_{\max}$$

As for $\sigma_{\min}$, for any fixed values of S'_t and S_t, there exists $r \in \mathbb{R}$ such that

$$S'_t = rS_t \tag{64}$$

Substituting into (62), we obtain

$$\frac{\sqrt{\sigma_1^2 + \sigma_2^2 r^2}}{1 + r}$$

Consider now the function $g(r)$ given by

$$g(r) = \frac{\sigma_1^2 + \sigma_2^2 r^2}{(1+r)^2}$$

Differentiating $g(r)$ with respect to r and setting $g'(r) = 0$, we find that $g(r)$ attains its minimum, say $r_{\min}$, when

$$r_{\min} = \frac{\sigma_1^2}{\sigma_2^2}$$

And thus, we obtain, under these conditions, that

$$\sigma_t \geqslant \frac{\sqrt{\sigma_1^2 + \sigma_2^2 r_{\min}^2}}{1 + r_{\min}} = \frac{\sigma_2 \sqrt{\sigma_1^2 \sigma_2^2 + \sigma_1^4}}{\sigma_1^2 + \sigma_2^2}$$
$$= \frac{\sigma_2 \sigma_1}{\sqrt{\sigma_1^2 + \sigma_2^2}}$$

Since, within our assumptions, we can always find $r \in \mathbb{R}$ for which (64) is satisfied, we obtain, as a consequence, explicit bounds for the initial price of the option, according to (63). In particular, by using strategy $\Delta_{\sigma_{\max}}$, that is, the hedging strategy associated with $\sigma_{\max}$, we are guaranteed to have at terminal time T at least $h(S_T)$, provided that we invest at least $\mathbb{E}\left[\frac{h(S_T)}{M_T}\right]$.

4 Optimal Hedging under Uncertainty

4.1 Initial Considerations

Having stated and proved Theorem (10), we thus have acquired a slight intuition on what to expect under misspecification of volatility. On Example (3.5), we have seen that, under the assumption of constant volatility on both assets, we were able to find (constant) lower and upper bounds for the stochastic volatility given by the resulting arithmetic mean of both assets. Theorem (10) then provided us with lower and upper bounds for the initial price of the option as well as what to expect with regard to the payout, thereby showing the robustness of the Black-Scholes Model.

This motivates us, therefore, to expand two notions. The first one, to define a random variable Y_t as our current Profits and Losses (hereafter $P\&L$) at time t, given by the difference between our hedging portfolio Π_θ (identified by the set of strategies $(\theta_t)_{0 \leqslant t \leqslant T}$) and the believed price of the contingent claim under the model of our choice (identified by $\bar{\sigma}$), $\bar{V}(t, S_t)$, with the further addition of some initial capital y_0

$$Y_t = y_0 + \int_0^t \theta_u dS_u - \bar{V}(t, S_t), \quad 0 \leqslant t \leqslant T \tag{65}$$

or, more plainly,

$$dY_t = \theta_u dS_u - d\bar{V}(t, S_t), \quad 0 \leqslant t \leqslant T. \tag{66}$$

In particular, Y_T is the payout at terminal T. And secondly with this definition, how to properly approach hedging in general, that is, firstly how to find an optimal set of strategies $\theta = (\theta_t)_{0 \leqslant t \leqslant T}$ under possible volatility misspecification and secondly, to this end, how to define "optimal".

Unlike Theorem (10), we are looking for a means to quantify the value of our set of strategies, with the further assumption that we are more or less certain that our estimate of $\bar{\sigma}$ is the true one, with regards to the true volatility governing stock S_t. For that, first notice that under model misspecification identified by $\bar{\sigma}$ and no interest rate, that is

$$dS_t = S_t \, \bar{\sigma}(t, S_t) dB_t \tag{67}$$

the price process $V(t,S_t)$ satisfies, accordingly, as in (58), the following

$$\frac{\partial}{\partial t}V(t,S_t)+\frac{1}{2}\bar{\sigma}_t^2S_t^2\frac{\partial^2}{\partial x^2}V(t,S_t)=0, \tag{68}$$

for all $0\leqslant t<T$, with

$$V(T,S_T)=h(S_T),$$

where $h(S_T)$ represents the payoff function of the vanilla option, as in Definition (15).

By defining the delta hedge

$$\Delta_{\bar{\sigma}}(t):=\frac{\partial}{\partial s}V(t,S_t),$$

to be the hedging strategy under no-arbitrage with respect to model (67), assuming $\bar{\sigma}=\sigma$ a.s, equation (66), by Itô's formula, becomes

$$\theta_t dS_t-\Delta_{\bar{\sigma}}(t)dS_t-\frac{\partial}{\partial t}V(t,S_t)dt-\frac{1}{2}\frac{\partial^2}{\partial x^2}V(t,S_t)\sigma_t^2S_t^2dt,$$

which, by (68), simplifies to

$$\theta_t dS_t-\Delta_{\bar{\sigma}}(t)dS_t+\frac{1}{2}\frac{\partial^2}{\partial x^2}V(t,S_t)\bar{\sigma}_t^2S_t^2dt-\frac{1}{2}\frac{\partial^2}{\partial x^2}V(t,S_t)\sigma_t^2S_t^2dt.$$

Hence, in integral form, this becomes

$$Y_t=y_0-V(0,S_0)+\int_0^t(\theta_u-\Delta_{\bar{\sigma}}(u))\,dS_u-\int_0^t\frac{1}{2}\frac{\partial^2}{\partial x^2}V(u,S_u)\,S_u^2(\sigma_u^2-\bar{\sigma}(u,S_u)^2)\,du,\quad 0\leqslant t\leqslant T.$$

Thus, we may ask ourselves what is the most sensible hedging strategy, that is, assuming that we fix *a priori* a model of our choice and given a subset of trading strategies, how can we maximize the following quantity

$$\max_\theta \mathbb{E}[U(Y_T)] \tag{69}$$

with U a concave function, the so called Utility function. This is a reasonable first step in defining an "optimal" set of strategies. Let us first confirm that the problem is well defined. For instance, if we assume that indeed $\bar{\sigma}(t,S_t)=\sigma_t$ and $\int_0^T(\theta_u-\Delta_{\bar{\sigma}}(u))dS_u$ is a supermartingale (both to rule out doubling strategies and to allow some friction on the market, though with no arbitrage opportunities), then, by Jensen´s inequality, (69) becomes

$$\mathbb{E}\left[U\Big(y_0-\bar{V}(0,S_0)+\int_0^T(\theta_u-\Delta_{\bar{\sigma}}(u))dS_u\Big)\right]\leqslant U\Big(y_0-\bar{V}(0,S_0)+\mathbb{E}\left[\int_0^T(\theta_u-\Delta_{\bar{\sigma}}(u))dS_u\right]\Big)\leqslant U(y_0-\bar{V}(0,S_0))$$

Thus, we obtain $\theta_u^*=\bar{\Delta}(u,S_u)$ as the best strategy pointwise, that is, precisely what one obtains in the case of a full replication of the underlying option as in (17). We have then the following

Lemma 8. *[9, Subsection 2.1, Lemma 2.1] Suppose $\sigma_t=\bar{\sigma}(t,S_t)$ and*

$$\bar{V}\in C^{1,2}((0,T)\times\mathbb{R}_+)\cap C([0,T]\times\mathbb{R}_+)$$

solves the PDE

$$\frac{\partial v_{\bar{\sigma}}}{\partial t}+\frac{1}{2}\bar{\sigma}(t,x)^2x^2\frac{\partial^2 v_{\bar{\sigma}}}{\partial x^2}=0,\quad t<T. \tag{70}$$

according to Lemma (7) *(with no interest rate). Then the delta hedge*

$$\bar{\Delta}_t = \frac{\partial}{\partial s}\bar{V}(t, S_t)$$

maximizes (69) *over any set* $\mathcal{A} \ni \bar{\Delta}$ *of strategies* θ *such that*

$$\int_0^T \left(\theta_t - \bar{\Delta}_t\right) \, dS_t$$

is a supermartingale.

Moreover,

$$\sup_{\theta \in \mathcal{A}} \mathbb{E}\left[U\left(x_0 + \int_0^T \theta_t \, dS_t - h(S_T)\right)\right] = U\big(x_0 - \bar{V}(0, s_0)\big) \, .$$

For the reason why we, from now on, assume a zero interest rate, we refer to [9, Remark 2.2]

4.2 General Setup and Introduction to the Penalized Optimization Problem

On this subsection, we follow directly [9, Subsection 2.2]. Fix a time horizon $T > 0$ and initial values $s_0 > 0$ and $y_0 \in \mathbb{R}$. Define

$$\Omega = \left\{\omega = (\omega_t^S, \omega_t^Y)_{t \in [0,T]} \in C([0,T]; \mathbb{R}^2) : \ \omega_0 = (s_0, y_0)\right\} .$$

That is, Ω is the space of all continuous $\mathbb{R}^2$-valued paths starting from (s_0, y_0), equipped with the topology of uniform convergence. Let $\mathcal{F} = \mathcal{B}(\Omega)$ denote the Borel σ-algebra generated by this topology. We introduce the canonical process $(S, Y) = ((S_t, Y_t))_{t \in [0,T]}$ on Ω by setting

$$S_t(\omega) = \omega_t^S, \qquad Y_t(\omega) = \omega_t^Y .$$

Thus, S and Y are simply the first and second coordinate processes, respectively. Finally, let $\mathbb{F} = (\mathcal{F}_t)_{t \in [0,T]}$ be the filtration generated by (S, Y), without imposing right-continuity or completion. Further, let $\bar{\mathcal{S}}$ be the collection of all triplets $(\theta, \sigma, \mathbb{Q})$ such that: $\theta = (\theta_t)_{t \in [0,T]}$ and $\sigma = (\sigma_t)_{t \in [0,T]}$ are progressively measurable processes, and $\mathbb{Q}$ is a probability measure on $(\Omega, \mathcal{F})$. Under $\mathbb{Q}$, the process S is a continuous local martingale whose quadratic variation satisfies

$$d[S, S]_t = S_t^2 \sigma_t^2 \, dt,$$

and the process Y is a continuous $\mathbb{Q}$ - semi martingale admitting a canonical decomposition

$$Y_t = y_0 - \bar{V}(0, S_0) + \int_0^t (\theta_u - \Delta_{\bar{\sigma}}(u)) dS_u - \int_0^t \frac{1}{2} \frac{\partial^2}{\partial x^2} \bar{V}(u, S_u) S_u^2 (\sigma_u^2 - \bar{\sigma}(u, S_u)^2), \quad 0 \leqslant t \leqslant T \quad (71)$$

Each pair of progressively measurable processes $\theta = (\theta_t)_{t \in [0,T]}$ and $\sigma = (\sigma_t)_{t \in [0,T]}$ determines a probability measure $\mathbb{Q}$ on $(\Omega, \mathcal{F})$ under which the canonical processes $(S_t)_{0 \leqslant t \leqslant T}$ and $(Y_t)_{0 \leqslant t \leqslant T}$ are defined.

Here, σ_t represents the true (or natural) volatility governing the dynamics of S_t, while θ_t denotes the hedging strategy implemented over the time interval $[0, T]$. Together, (θ, σ) generate the corresponding state processes (S, Y).

We now select a suitable subset

$$\mathcal{S} \subset \bar{\mathcal{S}},$$

and introduce sets of admissible controls

$$\mathcal{A} \quad \text{and} \quad \mathcal{V},$$

representing admissible hedging strategies and admissible volatility processes, respectively. These sets are chosen such that

$$\mathcal{A} \times \mathcal{V} \subset \mathcal{S},$$

and for every fixed strategy $(\theta_t)_{0\leqslant t\leqslant T} \in \mathcal{A}$ and every volatility process $(\sigma_t)_{0\leqslant t\leqslant T} \in \mathcal{V}$, the pair (θ, σ) induces a probability measure $\mathbb{Q}$ on $(\Omega, \mathcal{F})$ belonging to $\mathcal{S}$. Or more succinctly, denote $\mathcal{P}(\theta, \sigma)$ the set of probability measures induced by $\theta = (\theta_t)_{0\leqslant t\leqslant T}$ and $\sigma = (\sigma_t)_{0\leqslant t\leqslant T}$ (as the probability measure might not be unique). Then,

$$\mathcal{P}(\theta, \sigma) \neq \varnothing$$

It is important to point out that under the governing dynamics of S and Y given by

$$dS_t = S_t \sigma_t dB_t, \quad 0 \leqslant t \leqslant T$$

and

$$dY_t = (\theta_t - \Delta_{\bar{\sigma}}(t))dS_t - \frac{1}{2}\frac{\partial^2}{\partial x^2}\bar{V}(t, S_t)S_t^2(\sigma_t^2 - \bar{\sigma}(t, S_t)^2)dt, \quad 0 \leqslant t \leqslant T$$

with initial law given by $\delta_{(s_0, y_0)}$, if the weak solution satisfies uniqueness in the weak sense (12), then $\mathcal{P}(\theta, \sigma)$ is constituted by a single measure.

Now, introducing a penalty function $\alpha \in L^0(\Omega, \mathbb{F})$, the optimization problem that we are interested in is given by

$$\sup_{\theta \in \mathcal{A}} \inf_{\sigma \in \mathcal{V}} \inf_{\mathbb{Q} \in \mathcal{P}(\theta, \sigma)} \mathbb{E}_{\mathbb{Q}}[U(Y_T) + \alpha(\sigma)].$$

The penalty function $\alpha(\sigma)$ is of the form

$$\alpha(\sigma) = \frac{1}{\psi}\int_0^T U'(Y_t)\, f(t, Y_t, \sigma_t, S_t)\, dt,$$

where $\psi \in \mathbb{R}^+$ denotes the uncertainty parameter, and the function f is strictly convex with respect to σ_t. More precisely, for all (t, Y_t, S_t) fixed, the mapping

$$\sigma_t \mapsto f(t, Y_t, \sigma_t, S_t)$$

is strictly convex. Moreover, we further assume that

$$f(t, Y_t, \bar{\sigma}_t, S_t) = \frac{\partial}{\partial \sigma_t} f(t, Y_t, \bar{\sigma}_t, S_t) = 0.$$

The function f describes the shape of the penalty function, whereas ψ describes its magnitude: the lower the ψ, the more certain we are of our model. One good example of $f(t, Y_t, \sigma_t, S_t)$ is just the quadratic distance with respect to $\bar{\sigma}$. For the justification of the inclusion of $U'(Y_t)$ in the penalty function we refer to [9, Remark 2.6].

With this optimization problem, we have thus defined a good way of selecting an optimal set of strategies $(\theta_t)_{0\leqslant t\leqslant T}$ and further we defined what we mean by "optimal". Indeed, by essentially selecting the set of strategies that best perform against the worst possible true volatility or, more sensibly, against the worst adversarial true volatility. In turn, the weight given to this adversarial volatility is controlled by the penalty function in the above described manner.

Proceeding, let $\psi > 0$ be fixed. For any admissible strategy $\theta \in \mathcal{A}$, any admissible volatility process $\sigma \in \mathcal{V}$, and any probability measure $\mathbb{Q} \in \mathcal{P}(\theta, \sigma)$, we define the performance criterion of the hedging problem as

$$J^{\psi}(\sigma, \mathbb{Q}) := \mathbb{E}_{\mathbb{Q}}\left[U(Y_T) + \frac{1}{\psi}\int_0^T U'(Y_t)\, f(t, Y_t, \sigma_t, S_t)\, dt\right]$$

and the associated value function by

$$u(\psi) = u(\psi; \mathcal{A}, \mathcal{V}) := \sup_{\theta\in\mathcal{A}} \inf_{\sigma\in\mathcal{V}} \inf_{\mathbb{Q}\in\mathcal{P}(\theta,\sigma)} J^{\psi}(\sigma, \mathbb{Q}) \tag{72}$$

4.3 Assumptions and Main Theorem

Following directly [9, Subsection 3.1], we impose sufficient conditions that guarantee uniform boundedness of the relevant processes. In particular, we restrict attention to triplets $(\theta, \sigma, \mathbb{Q})$ such that the processes S, Y, and σ are $\mathbb{Q}$-almost surely uniformly bounded by a constant that does not depend on $(\theta, \sigma, \mathbb{Q})$.

Let constants

$$K > s_0 \vee s_0^{-1}, \qquad y_\ell < y_0 < y_u$$

be fixed. Define $\mathcal{S}$ as the collection of triplets $(\theta, \sigma, \mathbb{Q})$ satisfying

$$\sigma_t(\omega) \in [0, K], \qquad S_t(\omega) \in [K^{-1}, K], \qquad Y_t(\omega) \in (y_\ell, y_u), \tag{73}$$

for $(dt \times d\mathbb{Q})$-almost every $(t, \omega) \in [0, T] \times \Omega$.

We are interested in a Taylor expansion (with respect to ψ around 0) of the value function associated with the penalized control problem, $u(\psi)$, defined accordingly in (72).

To this end, we introduce a function $w^{\psi}(t, s, y)$, which at the terminal time T satisfies

$$w^{\psi}(T, s, y) = U(y).$$

We use w^{ψ} as a representation of the terminal utility $U(Y_T)$ evaluated along the state variables (S_t, Y_t), and we assume that for small ψ, it admits the expansion

$$w^{\psi}(t, s, y) = U(y) - U'(y)\bar{w}(t, s, y)\,\psi + U'(y)\tilde{w}(t, s, y)\,\psi^2 + o(\psi^2). \tag{74}$$

The functions $\bar{w}$ and $\tilde{w}$ correspond, therefore, to the first and second-order corrections in the expansion of $u(\psi)$, induced by the penalized control problem. They are characterized as solutions of two partial differential equations, which are given as follows.

Hypothesis 3. *Let*

$$D = (0,T) \times (K^{-1}, K) \times (y_\ell, y_u).$$

For every $y \in (y_\ell, y_u)$ we consider the linear second–order parabolic PDEs

$$\begin{aligned} \bar{w}_t(t,s,y) + \frac{1}{2}\bar{\sigma}(t,s)^2 s^2 \bar{w}_{ss}(t,s,y) + \bar{g}(t,s,y) &= 0, \quad (t,s) \in (0,T) \times (K^{-1},K), \\ \bar{w}(T,s,y) &= 0, \quad s \in (K^{-1},K), \\ \bar{w}(t,s,y) &= 0, \quad t \in (0,T),\ s \in \{K^{-1},K\}, \end{aligned} \tag{75}$$

and

$$\begin{aligned} \tilde{w}_t(t,s,y) + \frac{1}{2}\bar{\sigma}(t,s)^2 s^2 \tilde{w}_{ss}(t,s,y) + \tilde{g}(t,s,y) &= 0, \quad (t,s) \in (0,T) \times (K^{-1},K), \\ \tilde{w}(T,s,y) &= 0, \quad s \in (K^{-1},K), \\ \tilde{w}(t,s,y) &= 0, \quad t \in (0,T),\ s \in \{K^{-1},K\}. \end{aligned} \tag{76}$$

The source terms $\bar{g}$, $\tilde{g} : D \to \mathbb{R}$ are given by

$$\bar{g}(t,s,y) = \frac{(\bar{\sigma}(t,s)s^2\bar{V}_{ss}(t,s))^2}{2\, f''(t,s,y;\bar{\sigma}(t,s))}, \tag{77}$$

and

$$\begin{aligned} \tilde{g}(t,s,y) = {} & \frac{1}{6}\sigma(t,s,y)^3 f^{(3)}(t,s,y;\bar{\sigma}(t,s)) - \frac{1}{2}\sigma(t,s,y)^2 s^2 \bar{V}_{ss}(t,s) \\ & + \sigma(t,s,y)\bar{\sigma}(t,s)s^2\bar{V}_{ss}(t,s)\left(\bar{w}_y(t,s,y) - \bar{w}_{ss}(t,s,y)\right)\frac{U''(y)}{U'(y)} \\ & - \frac{1}{2}\bar{\sigma}(t,s)^2 s^2 \theta(t,s,y)^2 - \sigma(t,s,y)^2 f''(t,s,y;\bar{\sigma}(t,s))\, \bar{w}(t,s,y). \end{aligned} \tag{78}$$

The optimal trading strategy $\tilde{\theta}$ and the worst adversarial volatility $\tilde{\sigma}$ are defined by

$$\tilde{\theta}(t,s,y) = w_s(t,s,y) + \frac{U'(y)}{U''(y)}\, w_{sy}(t,s,y), \tag{79}$$

and

$$\tilde{\sigma}(t,s,y) = \frac{\bar{\sigma}(t,s)s^2\bar{V}_{ss}(t,s)}{f''(t,s,y;\bar{\sigma}(t,s))}. \tag{80}$$

These expressions will become at once clear in the next section, when we derive them explicitly in a heuristic setting. As to the conditions that allow for a such setting, we follow [9, Subsection 3.1, Assumption 3.2], hence the next hypothesis.

Hypothesis 4. *The following conditions hold.*

(i) PDE regularity. *There exist functions*

$$\bar{w}, \tilde{w} \in C^{1,2,2}(D) \cap C(D)$$

such that for each $y \in (y_\ell, y_u)$, $\bar{w}(\cdot,\cdot,y)$ and $\tilde{w}(\cdot,\cdot,y)$ are classical solutions of the PDEs above. Moreover,

$$|w_t|, |w_s|, |w_y|, |w_{ss}|, |w_{sy}|, |w_{yy}| \leqslant K, \qquad w \in \{\bar{w}, \tilde{w}\},$$

on D.

(ii) Reference local volatility. *The function*

$$\bar{\sigma} : [0, T] \times [K^{-1}, K] \to [0, K]$$

is Borel measurable and there exists $\varepsilon > 0$ such that

$$\varepsilon \leqslant \bar{\sigma}(t, s) \leqslant K - \varepsilon, \qquad (t, s) \in [0, T] \times (K^{-1}, K),$$

with boundary conditions

$$\bar{\sigma}(t, K) = \bar{\sigma}(t, K^{-1}) = 0$$

for all $t \in [0, T]$.

(iii) Reference value function. *The function*

$$\bar{V} : [0, T] \times [K^{-1}, K] \to \mathbb{R}$$

is Borel measurable and satisfies

$$\bar{V}(t, \cdot) \in C^2((K^{-1}, K)) \cap C([K^{-1}, K])$$

for all $t \in (0, T)$, together with the bound

$$|s^2 \bar{V}_{ss}(t, s)| \leqslant K, \qquad (t, s) \in (0, T) \times (K^{-1}, K).$$

(iv) Penalty function. *The function f is C^4 in ς. For $k = 2, 3, 4$ the derivatives $f^{(k)}$ satisfy*

$$\frac{1}{K} \leqslant f''(t, s, y; \varsigma) \leqslant K, \qquad |f^{(3)}(t, s, y; \varsigma)| \leqslant K, \qquad |f^{(4)}(t, s, y; \varsigma)| \leqslant K,$$

on $D \times [0, K]$.

(v) Utility function. *The utility function $U : \mathbb{R} \to \mathbb{R}$ is C^3 with*

$$U'(y) > 0, \qquad U''(y) < 0 \tag{81}$$

for all y.

Finally, since $\bar{\sigma}(t, K) = \bar{\sigma}(t, K^{-1}) = 0$, it is convenient to extend the definitions of σ, g, and $\bar{g}$ by setting

$$\sigma(t, s, y) = g(t, s, y) = \bar{g}(t, s, y) = 0,$$

for

$$t \in (0, T), \qquad s \in \{K^{-1}, K\}, \qquad y \in (y_\ell, y_u).$$

With this, we can state the following Theorem.

Theorem 11. *[9, Subsection 3.2, Theorem 3.4] Assume that Hypothesis* (4) *is satisfied and introduce the stopping time*

$$\tau := \inf \{t \in [0, T] : |Y_t - y_0| \geqslant 1\} \wedge T. \tag{82}$$

For every $\psi > 0$, define the perturbed strategy $\theta^\psi = (\theta_t^\psi)_{t\in[0,T]}$ and volatility $\sigma^\psi = (\sigma_t^\psi)_{t\in[0,T]}$ by

$$\theta_t^\psi = \Delta_t + \psi\, \tilde{\theta}(t, S_t, Y_t)\mathbf{1}_{\{t<\tau\}}, \tag{83}$$

$$\sigma_t^\psi = \bar{\sigma}(t, S_t) + \psi\, \tilde{\sigma}(t, S_t, Y_t), \tag{84}$$

where the functions $\tilde{\theta}$ and $\tilde{\sigma}$ are specified in (79) *and* (80)*, respectively.*

Let $\mathcal{A}$ and $\mathcal{V}$ be collections of progressively measurable processes such that, for all sufficiently small $\psi > 0$, we have

$$(\theta^\psi, \sigma^\psi) \in \mathcal{A} \times \mathcal{V} \subset \mathcal{S}.$$

Then, as $\psi \downarrow 0$, the value function admits the expansion

$$\begin{aligned} u(\psi) &= \sup_{\theta\in\mathcal{A}} \inf_{\sigma\in\mathcal{V}} \inf_{\mathbb{Q}\in\mathcal{P}(\theta,\sigma)} J^\psi(\sigma, \mathbb{Q}) && (85)\\ &= U(y_0) - U'(y_0)\, \bar{w}(0, s_0, y_0)\, \psi + U'(y_0)\, \tilde{w}(0, s_0, y_0)\, \psi^2 + o(\psi^2), && (86)\\ &= \inf_{\mathbb{Q}\in\mathcal{P}(\theta^\psi,\sigma^\psi)} J^\psi(\sigma^\psi, \mathbb{Q}) + o(\psi^2), \end{aligned}$$

where $\bar{w}$ and $\tilde{w}$ denote the solutions of PDEs (75) *and* (76)*, respectively.*

In particular, the strategy θ^ψ is optimal up to second order in ψ, that is, at order $O(\psi^2)$ within control $\mathcal{A}$. Likewise, σ^ψ represents a worst-case volatility up to order $O(\psi^2)$ within control $\mathcal{V}$.

The proof of the above Theorem can be found on [9, Subsection 5.1]. We will provide an heuristic proof in the next subsection, following [9, Subsection 4.1] . But, before that, some comments on Theorem (11).

According to [9, Subsection 3.2, Remark 3.5], it is not of small worth to take a look at the optimal controls (83) and (84). Indeed, if we observe more carefully the dynamics of the $P\&L$ described in (71) and further assume a linear perturbation on both controls as Theorem (11) does, it is easily seen, by performing a Taylor Expansion on the value function u (more concretely, to $U(Y_T)$), that a perturbation of the optimal strategy θ^ψ is only noticeable at order $O(\psi^2)$, since it only influences $P\&L$ through the diffusion term. By contrast, a perturbation of the optimal volatility σ^ψ directly affects the value function u at the leading order $O(\psi)$.

These observations help us explain the strong emphasis placed, in practice, on accurate model calibration, and, in particular, the potentially greater need of careful tuning local volatility models relative to adjusting our trading strategies. Indeed, this is as well reflected by the very definition of the stopping time τ in (82). In instances where a considerable profit or loss is observed, Theorem (11) dictates, from then on, the use of the unperturbed hedging strategy $\bar{\Delta}$ over any other perturbed strategy, with the clear goal of minimizing risk. This reliance on the reference model and its associated hedging strategy $\bar{\Delta}$, under such conditions, is partially justified by Theorem (10), which, as previously discussed, establishes the robustness of this approach.

In addition to these observations, it is also important to give an account on the nature of the value function itself, $w(t, s, y)$ described fully in Hypothesis (3). To that, we state the following Proposition, which gives an explicit representation of its linear term, $\bar{w}(t, s, y)$.

Proposition 6 (Feynman-Kac representation)**.** *Suppose Hypothesis* (4) *holds. Moreover, assume that $\bar{\sigma}(t, \cdot)$ is uniformly continuous on (K^{-1}, K) for each $t \in [0, T]$.*

Then for all $t \in [0,T]$, $s \in [K^{-1}, K]$, and $y \in (y_\ell, y_u)$,

$$\bar{w}(t,s,y) = \mathbb{E}^{t,s}_{\mathbb{Q}^{\bar{\sigma}}}\left[\int_t^T \frac{\big(\bar{\sigma}(u,S_u)S_u^2\bar{V}_{ss}(u,S_u)\big)^2}{2f''\big(u,S_u,y;\bar{\sigma}(u,S_u)\big)}\,du\right], \tag{87}$$

where the measure $\mathbb{Q}^{\bar{\sigma}}$ is such that governs dynamics

$$dS_u = S_u\bar{\sigma}(u,S_u)\,dW_u,$$

and expectation $\mathbb{E}^{t,s}_{\mathbb{Q}^{\bar{\sigma}}}[.]$ is taken accordingly with starting point t and $S_t = s$.

A proof of this Proposition can be found in [9, Subsection 5.2].

As found in [9, Subsection 3.2, Proposition 3.6], "we see that if f'' is constant, $\bar{w}$ measures the expected volatility-weighted cash-gamma accumulated over the remaining lifetime of the option." Thus, in instances where the stock S_t stays in high cash-gamma $(S_t^2\bar{V}_{ss})$ zones, $\tilde{w}$ will be high, with the incremented effect of $\bar{\sigma}$. In particular, in turbulent markets, we have an amplification of the value function. In particular, $\tilde{w}$ becomes large when the process S spends significant time in regions where the cash gamma, $S_t^2\bar{V}_{ss}(t,S_t)$, is high. This effect is further amplified by the reference volatility $\bar{\sigma}$. In particular, in more volatile market conditions, already high values of $\bar{w}$ are further amplified.

In addition, of the interplay between the optimal strategy θ^ψ and w^ψ, we refer to [9, Subsection 3.2, Discussion of the almost optimal strategy].

4.4 Heuristic Proof of Theorem (11)

Let $T > 0$, $\mathcal{A}$ and $\mathcal{V}$ be two admissible controls and, as before, for each pair (θ,σ), let $\mathbb{Q} := \mathbb{Q}^{\theta,\sigma}$ be the induced measure for such pair. Further, let $M_T^{\theta,\sigma}$ be an integrable process under $\mathbb{Q}^{\theta,\sigma}$.

Proposition 7 (Martingale Optimality Principle)**.** *Assume there exists $\sigma^* \in \mathcal{V}$ and $\theta^* \in$ A such that*

for each θ, M_T^{θ,σ^*} is a supermartingale

for each σ, $M_T^{\theta^*,\sigma}$ is a submartingale

holds. Then the following must hold

$$M_0 = \mathbb{E}[M_T^{\sigma^*,\theta^*}] = \sup_{\theta\in\mathcal{A}}\inf_{\sigma\in\mathcal{V}}\mathbb{E}[M_T^{\sigma,\theta}] = \inf_{\sigma\in\mathcal{V}}\sup_{\theta\in\mathcal{A}}\mathbb{E}[M_T^{\sigma,\theta}]$$

where the expectation is taken at each case under $P^{\sigma,\theta}$, for corresponding σ and θ. In particular, σ^* and θ^* are optimal controls.

Proof. First, notice that under $\mathbb{Q}^{\theta^*,\sigma^*}$, $M_T^{\theta^*,\sigma^*}$ is a Martingale. Therefore, we have immediately

$$M_0 = \mathbb{E}[M_T^{\sigma^*,\theta^*}]$$

In addition, we have

$$\sup_{\theta\in\mathcal{A}}\mathbb{E}[M_T^{\theta,\sigma^*}] \leqslant M_0 \leqslant \inf_{\sigma\in\mathcal{V}}\mathbb{E}[M_T^{\theta^*,\sigma}]$$

Thus, we have

$$M_0 \leqslant \inf_{\sigma\in\mathcal{V}} \mathbb{E}[M_T^{\theta^*,\sigma}] \leqslant \sup_{\theta\in\mathcal{A}} \inf_{\sigma\in\mathcal{V}} \mathbb{E}[M_T^{\theta,\sigma}] = \inf_{\sigma\in\mathcal{V}} \sup_{\theta\in\mathcal{A}} \mathbb{E}[M_T^{\theta,\sigma}] \leqslant \sup_{\theta\in\mathcal{A}} \mathbb{E}[M_T^{\theta,\sigma^*}] \leqslant M_0$$

Thus, the other equalities hold. □

Now, suppose that the dynamics of $M_t^{\theta,\sigma}$ are given by

$$dM_t^{\theta,\sigma} = dA_t^{\theta,\sigma} + u_t^{\theta,\sigma} dt$$

where $A_t^{\theta,\sigma}$ is a $\mathbb{Q}^{\theta,\sigma}$ - martingale. Then, clearly, for the pair (θ^*, σ^*), if we impose

(i) For each θ, M^{θ,σ^*} is a supermartingale,

(ii) For each σ, $M^{\theta^*,\sigma}$ is a submartingale.

by the Martingale Optimality Principle (7), we must have

$$u_t^{\theta^*,\sigma^*} = \sup_{\theta\in\mathcal{A}} \inf_{\sigma\in\mathcal{V}} u_t^{\theta,\sigma} = \inf_{\sigma\in\mathcal{V}} \sup_{\theta\in\mathcal{A}} u_t^{\theta,\sigma} = 0$$

with the value function u defined in (72).

Now, set $w := w(t,s,y)$ to be the value function, so that $w(T,s,y) = U(y)$ and $f(t, Y_t, \sigma_t, S_t) = f(\sigma_t)$. Thus, setting

$$M_t = w(t, S_t, Y_t) + \frac{1}{\psi}\int_0^t U'(Y_s) f(\sigma_s)\, ds$$

By Itô´s formula, we have

$$d(w(t,S_t,Y_t)) = w_t dt + w_s dS_t + w_y dY_t + w_{sy} d[S_t, Y_t] \tag{88}$$
$$+ \frac{1}{2} w_{ss} d[S_t, S_t] \frac{1}{2} w_{yy} d[Y_t, Y_t] \tag{89}$$

Let us assume, w.l.o.g, the following dynamics under $\mathbb{Q}^{\theta,\sigma}$

$$dS_t = S_t\sigma_t dB_t, \quad dY_t = b(t, S_t, \sigma_t)dt + (\theta_t - \bar{\Delta}_t) dS_t$$

where $b(t,s,\sigma)$ is defined so that $b(t,s,\bar{\sigma}) = 0$. As a consequence, we have the following

$$d[S,Y]_T = S_t^2\sigma_t^2(\theta_t - \bar{\Delta}_t)$$
$$d[S,S]_t = S_t^2\sigma_t^2$$
$$d[Y,Y]_t = S_t^2\sigma_t^2(\theta_t - \bar{\Delta}_t)^2$$

Thus, (88) becomes

$$A_t^{\theta,\sigma} dB_t + (w_t + w_y b(t,S,\sigma_t) + w_{sy}(\sigma_t - \bar{\Delta}_t) S_t^2\sigma_t^2 + \frac{1}{2} w_{ss} S_t^2\sigma_t^2 + w_{yy} S_t^2\sigma_t^2(\theta_t - \bar{\Delta}_t)^2) dt$$

According to the condition derived by the Martingale Optimality Principle (7), under the optimal pairs (θ^*, σ^*), we are left with the following equality, the *HJBI Equation*

$$w_t + \sup_{\theta\in\mathcal{A}} \inf_{\sigma\in\mathcal{V}} \big[\frac{1}{\psi} U'(y) f(\sigma_t) + w_y b(t,S,\sigma_t) + \frac{1}{2} S_t^2\sigma_t^2 [2w_{sy}(\theta_t - \bar{\Delta}_t) + w_{ss} + w_{yy}(\theta_t - \bar{\Delta}_t)^2]\big] = 0 \tag{90}$$

Now, taking into account Lemma (8), motivated by the fact that in the absence of model uncertainty the optimal strategy corresponds to the hedge $\bar{\Delta}$ with respect to the reference volatility $\bar{\sigma}$, we consider the reasonable ansatz

$$\begin{aligned}
\sigma^{\psi}(t,s,y) &= \bar{\sigma}(t,s) + \psi\tilde{\sigma}(t,s,y),\\
\theta^{\psi}(t,s,y) &= \bar{\Delta}(t,s) + \psi\tilde{\theta}(t,s,y),\\
w^{\psi}(t,s,y) &= U(y) - U'(y)\bar{w}(t,s,y)\psi + U'(y)\tilde{w}(t,s,y)\psi^2,
\end{aligned}$$

with the view to obtaining an asymptotic solution for small values of the uncertainty aversion parameter ψ. We are left to find $\bar{\sigma}$, $\tilde{\sigma}$, $\bar{\Delta}$, $\tilde{\theta}$, $\bar{w}(t,s,y)$ and $\tilde{w}(t,s,y)$. Notice that the form of the ansatz of w corresponds exactly to the expansion of the value function included in Theorem (11). Moreover, since $f(\bar{\sigma}) = f'(\bar{\sigma}) = 0$ and $b(\bar{\sigma}) = 0$ by assumption, we use the expansions

$$\begin{aligned}
f(\sigma) &\approx \frac{1}{2}f''(\bar{\sigma})(\sigma-\bar{\sigma})^2 + \frac{1}{6}f'''(\bar{\sigma})(\sigma-\bar{\sigma})^3,\\
b(\sigma) &= b'(\bar{\sigma})(\sigma-\bar{\sigma}) + \frac{1}{2}b''(\bar{\sigma})(\sigma-\bar{\sigma})^2
\end{aligned}$$

Remark 11. *On the above expansions, we stopped at the quadratic order, as the ansatz suggests, because we are interested in, at most, a quadratic behavior of the value function.*

From the ansatz we obtain

$$\begin{aligned}
w_t &= -U'(y)\bar{w}_t\psi + U'(y)\tilde{w}_t\psi^2,\\
w_y &= U'(y) - \big(U''(y)\bar{w} + U'(y)\bar{w}_y\big)\psi + \big(U''(y)\tilde{w} + U'(y)\tilde{w}_y\big)\psi^2,\\
w_{ss} &= -U'(y)\bar{w}_{ss}\psi + U'(y)\tilde{w}_{ss}\psi^2,\\
w_{sy} &= -\big(U''(y)\bar{w}_s + U'(y)\bar{w}_{sy}\big)\psi + \big(U''(y)\tilde{w}_s + U'(y)\tilde{w}_{sy}\big)\psi^2,\\
w_{yy} &= U''(y) - \big(U'''(y)\bar{w} + 2U''(y)\bar{w}_y + U'(y)\bar{w}_{yy}\big)\psi\\
&\quad + \big(U'''(y)\tilde{w} + 2U''(y)\tilde{w}_y + U'(y)\tilde{w}_{yy}\big)\psi^2.
\end{aligned}$$

Since

$$\sigma - \bar{\sigma} = \psi\tilde{\sigma},$$

we obtain

$$f(\sigma) = \frac{1}{2}f''(\bar{\sigma})\psi^2\tilde{\sigma}^2 + \frac{1}{6}f'''(\bar{\sigma})\psi^3\tilde{\sigma}^3,$$

and therefore

$$\frac{1}{\psi}U'(y)f(\sigma) = U'(y)\left[\frac{1}{2}f''(\bar{\sigma})\psi\tilde{\sigma}^2 + \frac{1}{6}f'''(\bar{\sigma})\psi^2\tilde{\sigma}^3\right].$$

Similarly,

$$b(\sigma) = b'(\bar{\sigma})\psi\tilde{\sigma} + \frac{1}{2}b''(\bar{\sigma})\psi^2\tilde{\sigma}^2,$$

so that

$$\begin{aligned}
w_y b(\sigma) &= \Big(U'(y) - (U''(y)\bar{w} + U'(y)\bar{w}_y)\psi\Big)b'(\bar{\sigma})\psi\tilde{\sigma} + U'(y)\frac{1}{2}b''(\bar{\sigma})\psi^2\tilde{\sigma}^2 + o(\psi^2)\\
&= U'(y)b'(\bar{\sigma})\psi\tilde{\sigma} - (U''(y)\bar{w} + U'(y)\bar{w}_y)b'(\bar{\sigma})\psi^2\tilde{\sigma} + U'(y)\frac{1}{2}b''(\bar{\sigma})\psi^2\tilde{\sigma}^2 + o(\psi^2).
\end{aligned}$$

Next, observe that

$$\sigma^2 = (\bar{\sigma} + \psi\tilde{\sigma})^2 = \bar{\sigma}^2 + 2\psi\bar{\sigma}\tilde{\sigma} + \psi^2\tilde{\sigma}^2, \tag{91}$$

and

$$\theta - \Delta_{\bar{\sigma}} = \psi\tilde{\theta}.$$

We now expand

$$2w_{sy}(\theta - \Delta_{\bar{\sigma}}) + w_{ss} + w_{yy}(\theta - \Delta_{\bar{\sigma}})^2.$$

Up to order ψ^2 this gives

$$\psi\Big[-U'(y)\bar{w}_{ss}\Big] + \psi^2\Big[-2(U''(y)\bar{w}_s + U'(y)\bar{w}_{sy})\tilde{\theta} + U'(y)\tilde{w}_{ss} + U''(y)\tilde{\theta}^2\Big]$$
$$+\, o(\psi^2)$$

Multiplying by $\frac{1}{2}S^2\sigma^2$, by expanding σ^2 as in (91), yields up to order $o(\psi^2)$

$$-\frac{1}{2}S^2\bar{\sigma}^2U'(y)\bar{w}_{ss}\psi + \frac{1}{2}S^2\bar{\sigma}^2\Big[-2(U''(y)\bar{w}_s + U'(y)\bar{w}_{sy})\tilde{\theta} + U'(y)\tilde{w}_{ss} + U''(y)\tilde{\theta}^2\Big]\psi^2$$
$$-\, S^2\bar{\sigma}\tilde{\sigma}U'(y)\bar{w}_{ss}\psi^2 + o(\psi^2).$$

Collecting all terms, the HJBI Equation (92) becomes

$$0 = -U'(y)\bar{w}_t\psi + U'(y)\tilde{w}_t\psi^2 \tag{92}$$
$$+ \sup_{\tilde{\theta}}\inf_{\tilde{\sigma}}\Bigg\{U'(y)\left[\frac{1}{2}f''(\bar{\sigma})\psi\tilde{\sigma}^2 + \frac{1}{6}f'''(\bar{\sigma})\psi^2\tilde{\sigma}^3\right] + U'(y)b'(\bar{\sigma})\psi\tilde{\sigma} \tag{93}$$
$$+\; U'(y)\frac{1}{2}b''(\bar{\sigma})\psi^2\tilde{\sigma}^2 - (U''(y)\bar{w} + U'(y)\bar{w}_y)b'(\bar{\sigma})\psi^2\tilde{\sigma} - \frac{1}{2}S^2\bar{\sigma}^2U'(y)\bar{w}_{ss}\psi \tag{94}$$
$$+ \frac{1}{2}S^2\bar{\sigma}^2\Big[-2(U''(y)\bar{w}_s + U'(y)\bar{w}_{sy})\tilde{\theta} + U'(y)\tilde{w}_{ss} + U''(y)\tilde{\theta}^2\Big]\psi^2 \tag{95}$$
$$-\, S^2\bar{\sigma}\tilde{\sigma}U'(y)\bar{w}_{ss}\psi^2 + o(\psi^2)\Bigg\}. \tag{96}$$

With the view towards determining explicitly $\tilde{\sigma}$, for $\psi << 1$, the linear order term dominates and thus we have to consider

$$-U'(y)\Big[\bar{w}_t - \frac{1}{2}f''(\bar{\sigma})\tilde{\sigma}^2 - b'(\bar{\sigma})\tilde{\sigma} + \frac{1}{2}S^2\bar{\sigma}^2\bar{w}_{ss}\Big]\psi. \tag{97}$$

Differentiating with respect to $\tilde{\sigma}$ and making expression (97) equal to 0, we thus obtain

$$\tilde{\sigma} = -\frac{b'(\bar{\sigma})}{f''(\bar{\sigma})}. \tag{98}$$

Indeed by looking at the second derivative, it is indeed an infimum. Applying the same reasoning to the quadratic order term, so that an optimal strategy may be obtained, we first group all the terms of interest together

$$\Big[U'(y)\tilde{w}_t + \frac{1}{6}U'(y)f'''(\bar{\sigma})\tilde{\sigma}^3 - (U''(y)\bar{w} + U'(y)\bar{w}_y)b'(\bar{\sigma})\tilde{\sigma} + U'(y)\frac{1}{2}b''(\bar{\sigma})\psi^2\tilde{\sigma}^2$$
$$+ \frac{1}{2}S^2\bar{\sigma}^2\Big[- 2(U''(y)\bar{w}_s + U'(y)\bar{w}_{sy})\tilde{\theta} + U'(y)\tilde{w}_{ss} + U''(y)\tilde{\theta}^2\Big]$$
$$- S^2\bar{\sigma}\tilde{\sigma}U'(y)\bar{w}_{ss}\Big]\psi^2. \tag{99}$$

Differentiating with respect to $\tilde{\theta}$ and making expression (99) equal to 0, it results in

$$S^2\bar{\sigma}^2\Big[- (U''(y)\bar{w}_s + U'(y)\bar{w}_{sy}) + U''(y)\tilde{\theta}\Big] = 0$$
$$\implies \tilde{\theta} = \bar{w}_s + \frac{U'(y)}{U''(y)}\bar{w}_{sy}. \tag{100}$$

Taking into account equation (92), we impose both expressions inside linear and quadratic orders to be equal to 0. As a consequence, by plugging the optimal controls (98) and (100) into (97) and (99), we end up, under the already determined optimal controls, with the following PDE´s pertaining to $\bar{w}$ and $\tilde{w}$.

$$U'(y)\Big[\bar{w}_t - \frac{1}{2}f''(\bar{\sigma})\frac{b'(\bar{\sigma})^2}{f''(\bar{\sigma})^2} + b'(\bar{\sigma})\frac{b'(\bar{\sigma})}{f''(\bar{\sigma})} + \frac{1}{2}S^2\bar{\sigma}^2\bar{w}_{ss}\Big] = 0$$
$$\implies \bar{w}_t + \frac{1}{2}S^2\bar{\sigma}^2\bar{w}_{ss} + \frac{b'(\bar{\sigma})^2}{2f''(\bar{\sigma})} = 0 \tag{101}$$

and

$$U'(y)\tilde{w}_t + U'(y)\frac{1}{6}f'''(\bar{\sigma})\tilde{\sigma}^3 - (U''(y)\bar{w} + U'(y)\bar{w}_y)b'(\bar{\sigma})\tilde{\sigma} \tag{102}$$
$$+ U'(y)\frac{1}{2}b''(\bar{\sigma})\tilde{\sigma}^2 + \frac{1}{2}S^2\bar{\sigma}^2\Big[- 2(U''(y)\bar{w}_s + U'(y)\bar{w}_{sy})\tilde{\theta} + U'(y)\tilde{w}_{ss} + U''(y)\tilde{\theta}^2\Big]$$
$$- S^2\bar{\sigma}\tilde{\sigma}U'(y)\bar{w}_{ss} = 0.$$

Now, from (100), the optimal control $\tilde{\theta}$, as a consequence, satisfies

$$\tilde{\theta}\,U''(y) = \bar{w}_sU''(y) + U'(y)\bar{w}_{sy}.$$

Thus, (102) simplifies to

$$U'(y)\tilde{w}_t + U'(y)\frac{1}{6}f'''(\bar{\sigma})\tilde{\sigma}^3 - (U''(y)\bar{w} + U'(y)\bar{w}_y)b'(\bar{\sigma})\tilde{\sigma}$$
$$+ U'(y)\frac{1}{2}b''(\bar{\sigma})\tilde{\sigma}^2 + \frac{1}{2}S^2\bar{\sigma}^2\Big[U'(y)\tilde{w}_{ss} - U''(y)\tilde{\theta}^2\Big]$$
$$- S^2\bar{\sigma}\tilde{\sigma}U'(y)\bar{w}_{ss} = 0,$$

which implies, in a compact form,

$$\tilde{w}_t + \frac{1}{2}S^2\bar{\sigma}^2U'(y)\tilde{w}_{ss} + \frac{1}{6}f'''(\bar{\sigma})\tilde{\sigma}^3 + \frac{1}{2}b''(\bar{\sigma})\tilde{\sigma}^2$$
$$- \bar{w}_yb'(\bar{\sigma})\tilde{\sigma} - S^2\bar{\sigma}\tilde{\sigma}\bar{w}_{ss}$$
$$- \frac{U''(y)}{U'(y)}\big[\bar{w}b'(\bar{\sigma})\tilde{\sigma} + \frac{1}{2}\tilde{\theta}^2S^2\bar{\sigma}^2\big] = 0. \tag{103}$$

In the case we are interested in, the drift term of Y_t is given by

$$b(t, S_t, \sigma_t) = \frac{1}{2} S_t^2 \bar{V}_{ss}(\bar{\sigma}_t^2 - \sigma_t^2)$$

which, in turn, implies

$$b'(\sigma_t) = -S_t^2 \bar{V}_{ss}\sigma_t, \quad b''(\sigma_t) = -S_t^2 \bar{V}_{ss}.$$

Thus, equations (101) and (103) become, respectively,

$$\bar{w}_t + \frac{1}{2} S^2 \bar{\sigma}^2 \bar{w}_{ss} + \frac{(S^2 \bar{V}_{ss}\bar{\sigma})^2}{2f''(\bar{\sigma})} = 0$$

and

$$\begin{aligned}
&\tilde{w}_t + \frac{1}{2} S^2 \bar{\sigma}^2 U'(y)\tilde{w}_{ss} + \frac{1}{6} f'''(\bar{\sigma})\tilde{\sigma}^3 - \frac{1}{2}\bar{\sigma}^2 S^2 \bar{V}_{ss} \\
&+ S^2 \bar{\sigma}\tilde{\sigma}(\bar{V}_{ss}\bar{w}_y - \bar{w}_{ss}) \\
&- \frac{U''(y)}{U'(y)}\big[- f''(\bar{\sigma})\tilde{\sigma}^2 \bar{w} + \frac{1}{2}\tilde{\theta}^2 S^2 \bar{\sigma}^2 \big] = 0
\end{aligned}$$

thereby obtaining PDEs described in (75) and (76). By the Martingale Optimality Principle (7), we have additionally

$$\mathbb{E}[M_T^{\sigma^*,\theta^*}] = M_0 = w(0, s_0, y_0)$$

which, by assumption, results in

$$U(y_0) - U'(y_0)\bar{w}(0, s_0, y_0)\psi + U'(y_0)\tilde{w}(t_0, s_0, y_0)\psi^2$$

which, as expected, coincides with (86).

5 Multiplicative Setting and Log-Utility Function

Theorem (11) provides us with a measure of the performance of our derived optimal trading strategy under the worst possible volatility, under Assumptions (4.3) (and through the value function w, described by PDEs (75) and (76), assuming that we are uncertain of our model, given by parameter ψ. We will look now into more detail on what particular form this collection of optimal pairs (θ_t^*, σ_t^*) assumes within a particular choice of a suitable Utility function. Unlike before, here we have a multiplicative approach and consider a logarithm utility function. This setting is therefore novel and not covered by [9]. For that, note first that we assumed that

$$dS_t = S_t \sigma_t dB_t$$

with σ_t as the true volatility governing asset S_t, though we incorrectly assume that it is governed by $\bar{\sigma}$, thus

$$\bar{V}_t + \frac{1}{2}\bar{\sigma}(t, S_t)^2 S_t^2 \bar{V}_{ss} = 0$$

Remark 12. *$\bar{V}_t$ and $\bar{V}_{ss}$ represent, respectively, $\frac{\partial \bar{V}(t,S_t)}{\partial t}$ and $\frac{\partial \bar{V}(t,S_t)^2}{\partial s^2}$.*

Suppose now that we define

$$\bar{S}_t = \frac{S_t}{\bar{V}(t, S_t)}, \tag{104}$$

and we are interested in trading the asset $\bar{S}_t$ instead of S_t. If we denote

$$X_t = x_0 + \int_0^t \theta_t \, dS_t,$$

as our wealth, taking into account our trading strategy provided by $(\theta_t)_{0 \leqslant t \leqslant T}$, we define accordingly

$$\bar{X}_t = \frac{X_t}{\bar{V}(t, S_t)}. \tag{105}$$

$\bar{X}_t$ serves as a geometric analogue of the P&L, Y_t, defined in (65). In this formulation, the condition $\tilde{X}_T = 1$ corresponds to replicating the option at maturity T. Naturally, if we denote by $\mathbb{Q}$ the measure associated with martingale S_t, we want to further find the martingale measure $\bar{\mathbb{Q}}$ associated with $\bar{S}_t$. First of all, let us make an analogue of Lemma (8), in order to find the analogue hedge strategy under these conditions on the call option applied to $\bar{S}_t$. To that, as in Lemma (8), if indeed $\sigma_t = \bar{\sigma}_t$, notice that, under measure $\bar{\mathbb{Q}}$, which satisfies

$$\frac{d\bar{\mathbb{Q}}}{d\mathbb{Q}} = \frac{\bar{V}(T, S_T)}{\bar{V}(0, S_0)}$$

we have,

$$\mathbb{E}_{\bar{\mathbb{Q}}}[\bar{S}_t|\mathcal{F}_s] = \frac{\bar{V}(0, S_0)}{\bar{V}(0, S_0)} \frac{\mathbb{E}_{\mathbb{Q}}[\bar{S}_t \bar{V}(T, S_T)|\mathcal{F}_s]}{\mathbb{E}_{\bar{\mathbb{Q}}}[\bar{V}(T, S_T)|\mathcal{F}_s]} = \frac{\mathbb{E}_{\mathbb{Q}}[\bar{S}_t \bar{V}(t, S_t)|\mathcal{F}_s]}{\bar{V}(s, S_s)}$$
$$= \frac{\mathbb{E}_{\mathbb{Q}}[S_t|\mathcal{F}_s]}{\bar{V}(s, S_s)} = \frac{S_s}{\bar{V}(s, S_s)} = \bar{S}_s$$

where the first equality is achieved through applying Bayes‘ rule applied to conditional expectations, the second through the tower property and the rest follow from the properties of the $\mathbb{Q}$-martingale S_t. Under $\bar{\mathbb{Q}}$ and any concave utility function U, we must have, by Jensen´s equality and by the martingale property of $\bar{S}_t$, under $\bar{\mathbb{Q}}$,

$$\mathbb{E}_{\bar{\mathbb{Q}}}\Big[U(\bar{X}_T)\Big] = \mathbb{E}_{\bar{\mathbb{Q}}}\Big[U(\bar{x}_0 + \int_0^T \theta_u \, d\bar{S}_u)\Big] \leqslant \mathbb{E}_{\bar{\mathbb{Q}}}\Big[U(\bar{x}_0)\Big]$$

which is attained at $\theta_t^* = 0$ a.s. Thus, if we define X_t^* to be

$$X_t^* = \bar{x}_0 \bar{V}(t, S_t),$$

the optimal geometric P&L is given by

$$\bar{X}_t^* = \frac{X_t^*}{\bar{V}(t, S_t)} = \bar{x}_0$$

and it follows that

$$\begin{aligned}
X_t^* &= \bar{x}_0 \bar{V}(t, S_t) = \bar{x}_0 \bar{V}(0, S_0) + \int_0^t \bar{x}_0 \bar{V}_s \, dS_u \\
&= \bar{x}_0 \bar{V}(0, S_0) + \int_0^t \bar{x}_0 \bar{V}(u, S_u) \frac{\bar{V}_s}{\bar{V}(u, S_u)} \, dS_u \\
&= X_0^* + \int_0^t X_u^* \frac{\bar{V}_s}{\bar{V}(u, S_u)} \, dS_u
\end{aligned}$$

By setting,

$$\pi_t = \frac{\bar{V}_s}{\bar{V}(t, S_t)} \tag{106}$$

we thus obtain

$$dX_t^* = X_t^* \pi_t dS_t, \quad d\bar{V}(t, S_t) = \bar{V}(t, S_t) \pi_t dS_t$$

and our hedging strategy (more precisely, the delta hedge adapted to this setting) becomes the proportion that we invest with respect to the stock we currently held. Thus, if we had before

$$X_t^\theta = x_0^\theta + \int_0^t \theta_u dS_u$$

we now reparametrize to

$$\begin{aligned}
X_t^\pi &= x_0^\pi + \int_0^t X_u^\pi \pi_u \, dS_u \\
\implies X_t^\pi &= x_0^\pi \exp\Big[\int_0^t \pi_u \, dS_u - \frac{1}{2} \int_0^t \pi_u^2 d[S, S]_u \Big]
\end{aligned}$$

If we now assume that

$$dS_t = S_t \sigma_t dB_t$$

then, we have by Itô

$$\begin{aligned}
d\bar{V}(t, S_t) &= \bar{V}_t \, dt + \bar{V}_s \, dS_t + \bar{V}_{ss} \sigma_t^2 S_t^2 \, dt \\
&= \bar{V}_s \, dS_t + \frac{1}{2} S_t^2 \bar{V}_{ss} \big(\sigma_t^2 - \bar{\sigma}(t, S_t)^2 \big) dt \\
&= \bar{V}(t, S_t) \left(\frac{\bar{V}_s}{\bar{V}(t, S_t)} \, dS_t + \frac{1}{2} S_t^2 \frac{\bar{V}_{ss}}{\bar{V}(t, S_t)} \big(\sigma_t^2 - \bar{\sigma}(t, S_t)^2 \big) dt \right) \\
\implies \bar{V}(t, S_t) &= \bar{V}(0, S_0) \exp\Bigg[\int_0^t \frac{\bar{V}_s}{\bar{V}(u, S_u)} \, dS_u \\
&\quad + \frac{1}{2} \int_0^t S_u^2 \frac{\bar{V}_{ss}}{\bar{V}(u, S_u)} \big(\bar{\sigma}(u, S_u)^2 - \sigma_u^2 \big) \, du \\
&\quad - \frac{1}{2} \int_0^t \frac{\bar{V}_s^2}{\bar{V}(u, S_u)^2} S_u^2 \sigma_u^2 \, du \Bigg]
\end{aligned}$$

As a consequence, we obtain

$$\bar{X}_t = \frac{X_t}{\bar{V}(t, S_t)} = \frac{\bar{x}_0}{\bar{V}(0, S_0)} \exp\Bigg[\int_0^t \left(\pi_u - \frac{\bar{V}_s}{\bar{V}(u, S_u)} \right) dS_u + \frac{1}{2} \int_0^t S_u^2 \frac{\bar{V}_{ss}}{\bar{V}(u, S_u)} \big(\bar{\sigma}(u, S_u)^2 - \sigma_u^2 \big) \, du$$

$$-\frac{1}{2}\int_0^t \left(\pi_u^2 - \frac{\bar{V}_s^2}{\bar{V}(u,S_u)^2}\right) S_u^2\sigma_u^2\, du\Bigg]$$

If indeed, we have $\bar{\sigma}_t = \sigma_t$ a.s, then clearly we have a full replication of the option and, as a result, $\bar{X}_T = 1$. Let us now find exactly the change of measure by which $\bar{S}_t$ becomes a $\bar{\mathbb{Q}}$-martingale. For that, assume that

$$\frac{d\bar{\mathbb{Q}}}{d\mathbb{Q}} = Z, \quad dZ_t = L_t dS_t \tag{107}$$

with $Z_T = Z$. First, notice that

$$\begin{aligned} d(Z_tS_t) &= S_t dZ_t + Z_t dS_t + L_t d[S,S]_t \\ &= (S_tL_t + Z_t)dS_t + L_tS_t^2\sigma_t^2 dt \end{aligned}$$

and so,

$$\begin{aligned} d\Big(\frac{Z_tS_t}{\bar{V}(t,S_t)}\Big) &= \frac{1}{\bar{V}(t,S_t)}d(Z_tS_t) - \frac{Z_tS_t}{\bar{V}(t,S_t)^2}d\bar{V}(t,S_t) - \frac{d[ZS,\bar{V}(t,S_t)]_t}{\bar{V}(t,S_t)^2} + \frac{Z_tS_t}{\bar{V}(t,S_t)^3}d[\bar{V}(t,S_t),\bar{V}(t,S_t)]_t \\ &= \Big[\frac{L_tS_t^2\sigma_t^2}{\bar{V}(t,S_t)} - \frac{Z_t}{\bar{V}(t,S_t)^2}\frac{1}{2}S_t^3\bar{V}_{ss}(\sigma_t^2 - \bar{\sigma}(t,S_t)^2) - \frac{(S_tL_t+Z_t)\bar{V}_sS_t^2\sigma_t^2}{\bar{V}(t,S_t)^2} + \frac{\bar{V}_s^2Z_tS_t^3\sigma_t^2}{\bar{V}(t,S_t)^3}\Big]dt + A_t dS_t \end{aligned}$$

with A_t some $\mathbb{Q}$- martingale. Thus, in order to make $\bar{S}_t$ a $\bar{\mathbb{Q}}$ martingale, $\frac{Z_tS_t}{\bar{V}(t,S_t)} = Z_t\bar{S}_t$ must be a $\mathbb{Q}$- martingale, provided that $\frac{d\bar{\mathbb{Q}}}{d\mathbb{Q}} = Z$. Thus, we must have

$$\frac{L_tS_t^2\sigma_t^2}{\bar{V}(t,S_t)} - \frac{Z_t}{\bar{V}(t,S_t)^2}\frac{1}{2}S_t^3\bar{V}_{ss}(\sigma_t^2 - \bar{\sigma}(t,S_t)^2) - \frac{(S_tL_t+Z_t)\bar{V}_sS_t^2\sigma_t^2}{\bar{V}(t,S_t)^2} + \frac{Z_tS_t^3\bar{V}_s^2\sigma_t^2}{\bar{V}(t,S_t)^3} = 0$$

And so, we derive L_t in the following way:

$$\frac{L_tS_t^2\sigma_t^2}{\bar{V}(t,S_t)} - \frac{Z_t}{\bar{V}(t,S_t)^2}\frac{1}{2}S_t^3\bar{V}_{ss}(\sigma_t^2 - \bar{\sigma}(t,S_t)^2) - \frac{(S_tL_t+Z_t)\bar{V}_sS_t^2\sigma_t^2}{\bar{V}(t,S_t)^2} + \frac{Z_tS_t^3\bar{V}_s^2\sigma_t^2}{\bar{V}(t,S_t)^3} = 0$$

$$L_t\bar{V}(t,S_t)^2 - \frac{1}{2}Z_tS_t\bar{V}_{ss}\Big(1 - \frac{\bar{\sigma}(t,S_t)^2}{\sigma_t^2}\Big)\bar{V}(t,S_t) - (S_tL_t+Z_t)\bar{V}_s\bar{V}(t,S_t) + Z_tS_t\bar{V}_s^2 = 0$$

$$L_t(\bar{V}(t,S_t)^2 - S_t\bar{V}_s\bar{V}(t,S_t)) = Z_t\bar{V}_s\bar{V}(t,S_t) - Z_tS_t\bar{V}_s^2 + \frac{1}{2}Z_tS_t\bar{V}_{ss}\Big(1 - \frac{\bar{\sigma}(t,S_t)^2}{\sigma_t^2}\Big)\bar{V}(t,S_t)$$

$$L_t = \frac{Z_t\bar{V}_s}{\bar{V}(t,S_t)} + \frac{1}{2}\frac{Z_tS_t\bar{V}_{ss}\Big(1 - \frac{\bar{\sigma}(t,S_t)^2}{\sigma_t^2}\Big)}{\bar{V}(t,S_t) - S_t\bar{V}_s}$$

$$L_t = \frac{Z_t\bar{V}_s}{\bar{V}(t,S_t)}\left[1 + \frac{1}{2}\frac{\bar{V}(t,S_t)S_t\bar{V}_{ss}}{\bar{V}_s(\bar{V}(t,S_t) - S_t\bar{V}_s)}\Big(1 - \frac{\bar{\sigma}(t,S_t)^2}{\sigma_t^2}\Big)\right]$$

With this particular choice of L_t, $\bar{S}_t$ is a $\bar{\mathbb{Q}}$-martingale through the change of measure $\frac{d\bar{\mathbb{Q}}}{d\mathbb{Q}} = Z_t$.

Now, take the logarithm as the Utility function. Note that this particular choice of the utility function does not satisfy condition (81). We consider the objective

$$\sup_{\pi} \mathbb{E}_{\bar{\mathbb{Q}}}\left[\log(\bar{X}_T)\right].$$

where the reference volatility $\bar{\sigma}(t, S_t)$ and the natural volatility σ_t are both fixed. Since we have

$$\begin{aligned}\bar{X}_t = \frac{X_t}{\bar{V}(t,S_t)} = \frac{\bar{x}_0}{\bar{V}(0,S_0)} \exp\Bigg[&\int_0^t \Big(\pi_u - \frac{\bar{V}_s}{\bar{V}(u,S_u)}\Big)\, dS_u \\ &+ \frac{1}{2}\int_0^t S_u^2 \frac{\bar{V}_{ss}}{\bar{V}(u,S_u)}\left(\bar{\sigma}_u^2 - \sigma_u^2\right) du \\ &- \frac{1}{2}\int_0^t \Big(\pi_u^2 - \frac{\bar{V}_s^2}{\bar{V}(u,S_u)^2}\Big) S_u^2\sigma_u^2\, du\Bigg]\end{aligned}$$

and we are maximizing over all possible strategies, we are interested in maximizing

$$\sup_{\pi} \mathbb{E}_{\bar{\mathbb{Q}}}\left[\int_0^T \pi_u dS_u - \frac{1}{2}\int_0^T \pi_u^2 S_u^2 \sigma_u^2 du\right]$$

Since $\frac{d\bar{\mathbb{Q}}}{d\mathbb{Q}} = Z$ and the dynamics of Z is given by (107), it is sufficient to solve under $\mathbb{Q}$ the following

$$\sup_{\pi} \mathbb{E}_{\mathbb{Q}}\left[\int_0^t \pi_u L_u d[S,S]_u - \frac{1}{2}\int_0^t \pi_u^2 S_u^2\sigma_u^2 Z_u du\right] = \mathbb{E}_{\mathbb{Q}}\left[\int_0^t S_u^2\sigma_u^2\Big[\pi_u L_u - \frac{1}{2}\pi_u^2 Z_u\Big]du\right] \tag{108}$$

Maximizing point-wise, by differentiating with respect to π_t, we obtain the optimal strategy π_t^*

$$\pi_t^* = \frac{L_t}{Z_t} = \frac{\bar{V}_s}{\bar{V}(t,S_t)}\Big[1 + \frac{1}{2}\frac{\bar{V}(t,S_t)S_t\bar{V}_{ss}}{\bar{V}_s(\bar{V}(t,S_t) - S_t\bar{V}_s)}\Big(1 - \frac{\bar{\sigma}(t,S_t)^2}{\sigma_t^2}\Big)\Big] \tag{109}$$

And, indeed, if $\sigma_t = \bar{\sigma}_t$ a.s, we obtain the hedging strategy obtained at (106).

One may further consider a more general formulation in which model uncertainty is incorporated through a penalty functional and we impose the most adversarial true volatility. We follow the same procedure as in previous subsection and try to approach the following problem

$$\sup_{\pi}\inf_{\sigma} E_{\bar{\mathbb{Q}}}\left[\log(\bar{X}_T) + \frac{1}{2\psi}\int_0^T a(\sigma_t)dt\right]$$

or, more explicitly,

$$\begin{aligned}\sup_{\pi}\inf_{\sigma} E_{\bar{\mathbb{Q}}}\Bigg[&\int_0^T \left(\pi_u - \frac{\bar{V}_s}{\bar{V}(u,S_u)}\right) dS_u \\ &+ \frac{1}{2}\int_0^T S_u^2 \frac{\bar{V}_{ss}}{\bar{V}(u,S_u)}\Big(\bar{\sigma}(u,S_u)^2 - \sigma_u^2\Big)\, du \\ &- \frac{1}{2}\int_0^T \left(\pi_u^2 - \frac{\bar{V}_s^2}{\bar{V}(u,S_u)^2}\right) S_u^2\sigma_u^2\, du \\ &+ \frac{1}{2\psi}\int_0^T a(\sigma_u)\, du\Bigg]\end{aligned}$$

If we perform the change of measure to $\mathbb{Q}$ and let $a(\sigma) = (\sigma_t - \bar{\sigma}(t, S_t))^2$, it results in

$$\sup_{\pi}\inf_{\sigma} E_{\mathbb{Q}}\left[\int_0^T\left(L_u S_u^2(\pi_u-\alpha_u)-\frac{1}{2}S_u^2 Z_u(\pi_u^2-\alpha_u^2)-\frac{1}{2}\beta_u Z_u\right)\sigma_u^2\,du+\frac{1}{2\psi}\int_0^T Z_u\big(\sigma_u-\bar{\sigma}(u,S_u)\big)^2\,du\right] \tag{110}$$

with

$$\alpha_u = \frac{\bar{V}_s}{\bar{V}(u, S_u)}, \quad \beta_u = S_u^2\frac{\bar{V}_{ss}}{\bar{V}(u, S_u)}$$

Before we proceed, we need to make sense of this optimization problem and further how to advance in a reasonable way. Contrary to the previous case, in which we approached optimization problem (108) under a fixed natural volatility, here we are faced with a different one, in which the measure $\mathbb{Q}$ varies by construction, with the result that a pointwise maximization is not allowed. We make this more explicit.

Suppose we have a random variable $f : \mathbb{R}^+ \times \mathbb{R} \to \mathbb{R}$ and we are interested in the problem

$$\sup_a \mathbb{E}[f(t, a)] \tag{111}$$

where the underlying measure is fixed and a is a random variable. In this case, we are allowed to treat f as a deterministic function and, as a consequence, to transform problem (111) into

$$\mathbb{E}[\sup_a f(t, a)] = \mathbb{E}[f(t, a^*)] \tag{112}$$

where a^* attains the supremum. Since the measure is fixed by hypothesis, if we choose , for each t, the corresponding a^* according to (112), the value of (111) corresponds then exactly to that attained by choosing a^* in (112). Thus, why the derivation of (109) was possible in the first place. If we, moreover, add one more layer of complexity and let $f : \mathbb{R}^+ \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$, while being interested in the optimization problem

$$\sup_a \inf_b \inf_{\beta\in\mathcal{P}(\mathbb{P}^b)} \mathbb{E}_{\mathbb{P}^b}[f(t, a, b)], \tag{113}$$

where now the underlying measure $\mathbb{P}^b$ depends directly on parameter b and is contained in the set of measures spanned by parameters a and b, the point-wise approach is not possible, since the attainable supremum, under a fixed measure, may not solve the optimization problem correctly. The only choice to overcome this is either finding a way to account for the choice of the current measure in some explicit way, or more simply, to assume that for a specific choice of b, say $\bar{b}$ and its associated measure $\mathbb{P}^{\bar{b}}$, that all measures $\mathbb{P}^b$ within a very small interval around $\bar{b}$, say $(\bar{b} - \epsilon, \bar{b} + \epsilon)$, for $\epsilon << 1$ behave as measure $\mathbb{P}^{\bar{b}}$. Thus, under this latter assumptions, optimization problem (113) would assume the form

$$\sup_a \inf_b \mathbb{E}_{\mathbb{P}^{\bar{b}}}[f(t, a, b)]$$

which, by applying Taylor Expansion reasonably on both parameters, does not give totally unreasonable results. We opted for this second approach and thus, we conjecture the following

Conjecture 1. *Assume that the candidate optimal volatility satisfies*

$$\sigma_t = \bar{\sigma}(t, S_t) + \mathcal{O}(\psi), \quad \textit{with } \psi \ll 1.$$

Then the corresponding probability measure $\mathbb{Q}^{\sigma}$ is close to $\mathbb{Q}^{\bar{\sigma}}$ in the sense that, for any admissible controls (π, σ), we have,

$$\mathbb{E}_{\mathbb{Q}^{\sigma}}\Big[\log(\bar{X}_T) + \frac{1}{2\psi}\int_0^T a(\sigma_t)dt\Big] = \mathbb{E}_{\mathbb{Q}^{\bar{\sigma}}}\Big[\log(\bar{X}_T) + \frac{1}{2\psi}\int_0^T a(\sigma_t)dt\Big] + \mathcal{O}(\psi).$$

Consequently, up to an error of order $\mathcal{O}(\psi)$, the optimization problem

$$\sup_{\pi}\inf_{\sigma}\mathbb{E}_{\mathbb{Q}^{\sigma}}[\cdot]$$

can be approximated by

$$\sup_{\pi}\inf_{\sigma}\mathbb{E}_{\mathbb{Q}^{\bar{\sigma}}}[\cdot],$$

i.e., the optimization can be carried out under the fixed reference measure $\mathbb{Q}^{\bar{\sigma}}$.

Assuming Conjecture (1), the optimization problem (110) transforms then into

$$\sup_{\pi}\inf_{\sigma} E_{\mathbb{Q}^{\bar{\sigma}}}\left[\int_0^T\left(L_u S_u^2(\pi_u-\alpha_u)-\frac{1}{2}S_u^2 Z_u(\pi_u^2-\alpha_u^2)-\frac{1}{2}\beta_u Z_u\right)\sigma_u^2\,du+\frac{1}{2\psi}\int_0^T Z_u\big(\sigma_u-\bar{\sigma}(u,S_u)\big)^2\,du\right]+\mathcal{O}(\psi)$$

where $\mathbb{Q}^{\bar{\sigma}}$ denotes the measure under dynamics

$$dS_t = S_t\bar{\sigma}(t,S_t)dB_t$$

Similarly to the formulation above, we now need to come up with reasonable ansatz for both controls under this formulation. With the analogue of Lemma (8) together with our prior estimate of the true volatility given by the local volatility model, it is reasonable to assume the following ansatz, with $\psi << 1$

$$\sigma^{\psi}(t,s,y) = \bar{\sigma}(t,s) + \psi\,\tilde{\sigma}(t,s,y) \tag{114}$$
$$\pi(t,s,y) = \alpha(t,s) + \psi\,\tilde{\pi}(t,s,y) \tag{115}$$

The difficulty now resides in grouping all terms according to their respective order with regard to ψ, followed by finding the worst volatility in the linear order and the optimal strategy in the quadratic order. Before we actually proceed, notice that

$$L_t = \frac{Z_t\bar{V}_s}{\bar{V}(t,S_t)}\Big[1+\frac{1}{2}\frac{\bar{V}(t,S_t)S_t\bar{V}_{ss}}{\bar{V}_s(\bar{V}(t,S_t)-S_t\bar{V}_s)}\Big(1-\frac{\bar{\sigma}(t,S_t)^2}{\sigma_t^2}\Big)\Big]$$

thus containing a control we are interested in. Set, to that end,

$$c(S_t,\bar{V}(t,S_t),Z_t) = Z_t\frac{\bar{V}_s}{\bar{V}(t,S_t)}\Big[1+\frac{1}{2}\frac{\bar{V}(t,S_t)S_t\bar{V}_{ss}}{\bar{V}_s(\bar{V}(t,S_t)-S_t\bar{V}_s)}\Big]$$

and

$$d(S_t,\bar{V}(t,S_t),Z_t) = -\frac{1}{2}\frac{Z_tS_t\bar{V}_{ss}}{(\bar{V}(t,S_t)-S_t\bar{V}_s)}$$

To ease notation, say $c(S_t,\bar{V}(t,S_t),Z_t) = c_t$, $d(S_t,\bar{V}(t,S_t),Z_t) = d_t$ and $\bar{\sigma}(t,S_t) = \bar{\sigma}$. Then, L_t can be written as

$$c_t + \frac{d_t}{\sigma_t^2} \tag{116}$$

Plugging in the ansatz (114) into (116), we obtain

$$
\begin{aligned}
c_t + \frac{d_t}{(\bar{\sigma}_t + \psi\tilde{\sigma}_t)^2} &= c_t + \frac{d_t}{\bar{\sigma}_t{}^2 + 2\bar{\sigma}_t\tilde{\sigma}_t\psi + \tilde{\sigma}_t{}^2\psi^2} \\
&= c_t + \frac{d_t}{\bar{\sigma}_t{}^2} \cdot \frac{1}{1 + \frac{2\tilde{\sigma}_t}{\bar{\sigma}_t}\psi + \frac{\tilde{\sigma}_t{}^2}{\bar{\sigma_T}{}^2}\psi^2} \\
&= c_t + \frac{d_t}{\bar{\sigma}_t{}^2}\left[1 - \left(\frac{2\tilde{\sigma}_t}{\bar{\sigma}_t}\psi + \frac{\tilde{\sigma}_t{}^2}{\bar{\sigma}_t{}^2}\psi^2\right) + \left(\frac{2\tilde{\sigma}_t}{\bar{\sigma}_t}\psi\right)^2 + o(\psi^2)\right] \\
&= c_t + \frac{d_t}{\bar{\sigma}_t{}^2}\left[1 - \frac{2\tilde{\sigma}_t}{\bar{\sigma}_t}\psi - \frac{\tilde{\sigma}_t{}^2}{\bar{\sigma}_t{}^2}\psi^2 + \frac{4\tilde{\sigma}_t{}^2}{\bar{\sigma}_t{}^2}\psi^2 + o(\psi^2)\right] \\
&= c_t + \frac{d_t}{\bar{\sigma}^2} - \frac{2d_t\tilde{\sigma}}{\bar{\sigma}^3}\psi + \frac{3d_t\tilde{\sigma}^2}{\bar{\sigma}^4}\psi^2 + o(\psi^2)
\end{aligned}
$$

Plugging both (115) and (114) into (110), and likewise easing notation by letting $\tilde{\sigma}(t,s,y) = \tilde{\sigma}_t$ and $\tilde{\pi}(t,s,y) = \tilde{\pi}_t$, we thus obtain

$$
\begin{aligned}
&\Bigg[\Bigg(c_t + \frac{d_t}{\bar{\sigma}_t{}^2} - \frac{2d_t\tilde{\sigma}_t}{\bar{\sigma}_t{}^3}\psi + \frac{3d_t\tilde{\sigma}_t{}^2}{\bar{\sigma}_t{}^4}\psi^2 + o(\psi^2)\Bigg)S_t^2\psi\tilde{\pi}_t\Bigg) \\
&- \frac{1}{2}S_t^2 Z_t[2\psi\tilde{\pi}_t\alpha_t + \psi^2\tilde{\pi}_t^2] - \frac{1}{2}\beta_t Z_t\Bigg](\bar{\sigma}_t + \psi\tilde{\sigma}_t)^2 \\
&+ \frac{1}{2}Z_t\psi\tilde{\sigma}_t^2
\end{aligned}
$$

Grouping all the terms according to their respective order, we obtain

$$
\begin{aligned}
&- \frac{1}{2}\beta_t Z_t\bar{\sigma}_t^2 \\
&+ \left[\left(c_t + \frac{d_t}{\bar{\sigma}_t^2}\right)S_t^2\tilde{\pi}_t\bar{\sigma}_t^2 - S_t^2 Z_t\tilde{\pi}_t\alpha_t\bar{\sigma}_t^2 - \beta_t Z_t\bar{\sigma}_t\tilde{\sigma}_t + \frac{1}{2}Z_t\tilde{\sigma}_t^2\right]\psi \\
&+ \left[2\bar{\sigma}_t\tilde{\sigma}_t S_t^2\tilde{\pi}_t\left(c_t + \frac{d_t}{\bar{\sigma}_t^2}\right) - \frac{1}{2}S_t^2 Z_t\tilde{\pi}_t^2\bar{\sigma}_t^2 - 2\bar{\sigma}_t\tilde{\sigma}_t S_t^2 Z_t\tilde{\pi}_t\alpha_t - \frac{1}{2}\beta_t Z_t\tilde{\sigma}_t^2 - \frac{2d_t\tilde{\sigma}_t}{\bar{\sigma}_t}S_t^2\tilde{\pi}_t\right]\psi^2 + o(\psi^2)
\end{aligned}
$$

which simplifies to

$$
\begin{aligned}
&- \frac{1}{2}\beta_t Z_t\bar{\sigma}_t^2 \\
&+ \left[\left(c_t + \frac{d_t}{\bar{\sigma}_t^2}\right)S_t^2\tilde{\pi}_t\bar{\sigma}_t^2 - S_t^2 Z_t\tilde{\pi}_t\alpha_t\bar{\sigma}_t^2 - \beta_t Z_t\bar{\sigma}_t\tilde{\sigma}_t + \frac{1}{2}\tilde{\sigma}_t^2\right]\psi \\
&+ \left[2\bar{\sigma}_t\tilde{\sigma}_t S_t^2\tilde{\pi}_t c_t - \frac{1}{2}S_t^2 Z_t\tilde{\pi}_t^2\bar{\sigma}_t^2 - 2\bar{\sigma}_t\tilde{\sigma}_t S_t^2 Z_t\tilde{\pi}_t\alpha_t - \frac{1}{2}\beta_t Z_t\tilde{\sigma}_t^2\right]\psi^2 + o(\psi^2)
\end{aligned}
$$

Using the same procedure, we first analyze the linear term in ψ. Minimizing with respect to $\tilde{\sigma}_t$, yields the first-order condition

$$-\beta_t Z_t\bar{\sigma}_t + Z_t\tilde{\sigma}_t = 0,$$

and hence

$$\tilde{\sigma}_t^* = \beta_t\bar{\sigma}_t.$$

Indeed, since the second derivative is positive, this critical point corresponds to a minimum.

Next, we consider the quadratic term. Differentiating the quadratic coefficient with respect to $\tilde{\pi}_t$ gives

$$2\bar{\sigma}_t\tilde{\sigma}_t^* S_t^2 c_t - S_t^2 Z_t \bar{\sigma}_t^2 \tilde{\pi}_t - 2\bar{\sigma}_t\tilde{\sigma}_t^* S_t^2 Z_t \alpha_t.$$

Solving for the optimal control yields

$$\tilde{\pi}_t^* = \frac{2\bar{\sigma}_t\tilde{\sigma}_t^*(c_t - Z_t\alpha_t)}{Z_t\bar{\sigma}_t^2}.$$

Further, substituting $\tilde{\sigma}_t^* = \beta_t\bar{\sigma}_t$ gives

$$\tilde{\pi}_t^* = \frac{2\beta_t(c_t - Z_t\alpha_t)}{Z_t}.$$

All in all, plugging everything back, we obtain the following optimal controls

$$\sigma_t^* = \bar{\sigma}_t + \psi\bar{\sigma}_t S_t^2 \frac{\bar{V}_{ss}}{\bar{V}(t, S_t)} \tag{117}$$

$$\begin{aligned}\pi_t^* &= \frac{\bar{V}_s}{\bar{V}(t,S_t)} + \psi 2\beta_t\Big[\frac{\bar{V}_s}{\bar{V}(t,S_t)}\Big\{1 + \frac{1}{2}\frac{\bar{V}(t,S_t)S_t\bar{V}_{ss}}{\bar{V}_s(\bar{V}(t,S_t) - S_t\bar{V}_s)}\Big\} - \frac{\bar{V}_s}{\bar{V}(t,S_t)}\Big] \\ &= \frac{\bar{V}_s}{\bar{V}(t,S_t)} + \psi\frac{\beta_t S_t\bar{V}_{ss}}{(\bar{V}(t,S_t) - S_t\bar{V}_s)} = \frac{\bar{V}_s}{\bar{V}(t,S_t)} + \psi\frac{S_t^3\bar{V}_{ss}^2}{\bar{V}(t,S_t)(\bar{V}(t,S_t) - S_t\bar{V}_s)}\end{aligned} \tag{118}$$

The previous derivations provide an explicit asymptotic characterization of the robust optimal investment problem under multiplicative dynamics and logarithmic utility. In particular, we derive explicit first-order representations for the worst-case volatility, given by (117), quantifying the deviation from the reference local volatility model, together with the associated optimal hedging strategy, given by (118), which can be interpreted as a first-order correction to the classical delta hedge associated with this multiplicative setting.

Hence, similarly to the optimal controls we have constructed in the previous section, this robust optimization problem admits, for small uncertainty aversion parameter ψ, an explicit perturbative structure around the reference local volatility model identified with $\bar{\sigma}$. In particular, the leading-order terms recover the unperturbed local volatility model and its associated hedging strategy, while the linear corrections quantify the investor's adjustment against model misspecification and volatility uncertainty.

Had we more time, we would have aimed not only to investigate more thoroughly the validity of Conjecture (1) and to extend the multiplicative setting to more general utility functions, but also to study the practical implications of the model through numerical simulations. In particular, this would allow us to compare the performance of the derived optimal hedging strategy against other hedging approaches under different market scenarios.

References

[1] L.C.G. Rogers and David Williams, *Diffusions, Markov Processes and Martingales*, Cambridge University Press (2000)

[2] Xi Geng, *Stochastic Calculus Lecture Notes* Carnegie Mellon University

[3] Mathias Beiglböck, *Mathematical Finance Lecture Notes*

[4] Nobuyuki Ikeda and Shinzo Watanabe, *Stochastical Differential Equations and Diffusion Processes*, North-Holland Publishing Company (1981)

[5] S.L. Sobolev, *Applications of functional analysis in mathematical physics*, Amer. Math. Soc., Providence, 1963.

[6] El Karoui, N., Jeanblanc-Pique, M., Shreve, S.E.Math. Finance. 8(2), 93–126 (1998): *Robustness of the Blacks-Scholes Formula*

[7] Bergman, Y. Z., B. D. Grundy, and Z. Wiener, *General Properties of Option Prices*, 1996, J. Finance 51, 1573–1610.

[8] Kunita, *Stochatic Flows and Stochastic Differential Equations*, Cambridge, UK: Cambridge University Press, 1990

[9] Sebastian Herrmann, Frank Thomas Seifried and Johannes Muhle-Karbe. Finance Stoch (2017) 21:1–64: *Hedging with Small Uncertainity Aversion*

[10] Hans Föllmer and Alexander Schied, *Stochastic Finance*, 2004 Walter de Gruyter GmbH & Co. KG, 10785 Berlin, Germany

[11] D. Lamberton and B. Lapeyre, *Introduction to Stochastic Calculus Applied to Finance*, 1996 by Chapman & Hall

[12] Steven. E. Shreve, *Stochastic Calculus for Finance II: Continuous-Time Models*, 2004 Springer Science + Business Media, Inc

[13] Sühan Altay and Uwe Schmock, *Lecture Notes on the Yamada–Watanabe condition for the pathwise uniqueness of solutions of certain Stochastic Differential Equations*

[14] Toshio Yamada and Shinzo Watanabe, *On the uniqueness of solutions of stochastic differential equations* J. Math. Kyoto Univ. (JMKYAZ) 11-1 (1971) 155-167

[15] Tomas Björk, *Arbitrage Theory in Continuous Time* Oxford University Press, 4th Edition, 2020