Portfolio Representation as Applied to Model Points Selection for ALM in Life Insurance

Autor: Enrico Ferri

Tese de doutoramento UDC / Ano 2018

Director/a: Pilar Barrios Gómez, Carlos Vázquez Cendón

Programa de doutoramento en Métodos Matemáticos e Simulación Numérica en Enxeñaría e Ciencias Aplicadas

UNIVERSIDADE DA CORUÑA

UNIVERSIDADE DA CORUÑA

PH.D. THESIS

Portfolio Representation as Applied to Model Points Selection for ALM in Life Insurance

AUTHOR:

Enrico Ferri

SUPERVISORS:

Pilar Barrios Gómez

Carlos Vázquez Cendón

TESE PRESENTADA PARA A OBTENCIÓN DO TÍTULO DE DOUTOR NA
UNIVERSIDADE DA CORUÑA
DEPARTAMENTO DE MATEMÁTICAS
FACULTADE DE INFORMÁTICA, A CORUÑA (SPAIN)
OUTUBRO, 2018

UNIVERSIDADE DA CORUÑA

Los de abajo firmantes hacen constar que son los directores de la Tesis Doctoral titulada **"Portfolio Representation as Applied to Model Points Selection for ALM in Life Insurance"** desarrollada por Enrico Ferri, cuya firma también se incluye, dentro del programa de doctorado **"Métodos Matemáticos y Simulación Numérica en Ingeniería y Ciencias Aplicadas"** en el Departamento de Matemáticas (Universidade da Coruña), dando su consentimiento para su presentación y posterior defensa.

The undersigned hereby certify that they are the supervisors of the Thesis entitled **"Portfolio Representation as Applied to Model Points Selection for ALM in Life Insurance"** developed by Enrico Ferri, whose signature is also included, inside the Ph.D Program **"Mathematical Modelling and Numerical Methods in Applied Sciences and Engineering"** at the Department of Mathematics (University of A Coruña), consenting to its presentation and posterior defense.

October, 2018

DIRECTORES:

Pilar Barrios Gómez

Carlos Vázquez Cendón

DOCTORANDO:

Enrico Ferri

Funding

This research has been funded by the following project:

MSCA-ITN-EID - European Industrial Doctorates (WAKEUPCALL Grant Agreement 643045).

Funded under: H2020–EU.1.3.1. – Fostering new skills by means of excellent initial training of researchers.

To my parents Piera and Silverio

Acknowledgements

I owe my great appreciation to my advisors Pilar Barrios, partner at Afi, and Carlos Vázquez, professor at University of A Coruña, for supervising and coordinating the development of this thesis over these years. Their valuable supervision has greatly improved the quality of this work.

Besides my advisors, I should like to express my deepest and sincere gratitude to José Luis Fernández, professor at Autonomous University of Madrid, whose outstanding guidance has been crucial for carrying out this thesis successfully. His suggestions as well as his constructive criticism constantly led and enhanced my research over the last years. His willingness to share time and experience with me has been invaluable.

A special thanks goes also to Ana María Ferreiro and José Antonio García, professors as University of A Coruña, for their collaboration and their help related to the computational issues of this work. The opportunity to work with them turned out to be strongly constructive for realizing this thesis.

I gratefully acknowledge the financial sources received under the auspices of the Wakeupcall project, funded in the framework of the EU Horizon2020 programme for research and innovation. In this regards, I would like also to thank Pasquale Cirillo, professor at Delft University of Technology, Cornelis W. Oosterlee, professor at Delft University of Technology and CWI in Amsterdam, and Andrea Pascucci, professor at University of Bologna, for coordinating the project jointly with professor Carlos

Vázquez. I thank them for providing me with the opportunity to give my contribution within this project. Moreover, my appreciation goes also to Anastasia Borovykh, Ki-Wai Chau, Andrea Fontanari, Sidy Diop and Zuzana Krajčovičová, who participated with me as early stage researchers in this project, for the many enjoyable and inspiring shared moments.

I thank all of the faculty and administrative staff at University of A Coruña, Afi and Autonomous University of Madrid for their help throughout the years. A special thanks goes also to my fellow PhD students Alberto Dayan, Alejandro Mas, Fran Mengual, Pablo Monetero and Beatriz Salvador.

I wish to give my deepest thanks to my girlfriend Paloma, for her unwavering motivation and affection. She constantly stood beside me throughout these years, sharing both failures and successes.

Finally, I would like to express my heartfelt gratefulness to my parents Piera and Silverio, for believing in me at any time. They gave me moral support and constant encouragement throughout the entire development of the present work. This thesis is dedicated to them.

Table of Contents

Abstract

The main goal of this work is represented by the study of the constrained portfolio representation problem. It consists in defining the dynamics of those portfolios within a certain class that best represents the inherent risk structure of a given financial exposure. An interesting example of this question arises in the context of ALM for life insurance firms, where the scale of a given portfolio of policies is to be reduced for analysis and management purposes.

This problem is initially formalized and addressed on a general setting, within the theory of stochastic integration in UMD Banach spaces and the Malliavin calculus developed in such a framework. Next, a numerical study applied to portfolios of term insurance policies is discussed when considering a LIBOR Market Model describing the time evolution of the forward rates term structure.

On the other hand, the study of the asymptotic stability of a given class of risk estimators, when enlarging the informative data set, provides a further goal of the present work. In this respect, a refined notion of robustness which applies when dealing with random probability measures is characterized in terms of the convergence properties displayed by the empirical process associated to a given sequence of stationary random variables.

Resumen

El principal objetivo de este trabajo viene dado por el estudio del problema de representación de carteras con restricciones. Consiste en definir las dinámicas de aquellas carteras que dentro de una cierta clase mejor representan la estructura de los riesgos inherentes de una exposición financiera dada. Un ejemplo interesante de esta cuestión surge en el contexto del ALM para compañías de seguros de vida, donde el tamaño de una cartera de pólizas debe ser reducido para su análisis y gestión.

Inicialmente, este problema se formaliza y aborda en un contexto general, dentro de la teoría de integración estocástica en espacios de Banach UMD y del cálculo de Malliavin desarrollado en ese marco. A continuación, se analiza un estudio numérico aplicado a carteras de pólizas de seguros a plazo, cuando se considera un LIBOR Market Model para describir la evolución temporal de los tipos de interés forward.

Por otro lado, el estudio de la estabilidad asintótica de una clase dada de estimadores del riesgo, cuando se amplía el conjunto de datos, proporciona un objetivo adicional al presente trabajo. En este sentido, una noción refinada de robustez, que se aplica cuando se manejan medidas de probabilidad aleatorias, se caracteriza en términos de las propiedades de convergencia verificadas por los procesos empíricos asociados a una sucesión dada de variables aleatorias estacionarias.

Resumo

O principal obxectivo deste traballo vén dado polo estudo do problema de representación de carteiras con restricións. Consiste en definir as dinámicas daquelas carteiras que dentro dunha certa clase mellor representan a estrutura dos riscos inherentes dunha exposición financeira dada. Un exemplo interesante desta cuestión xorde no contexto do ALM para compañías de seguros de vida, onde o tamaño dunha carteira de pólizas debe ser reducido para a súa análise e xestión.

Inicialmente, este problema formalízase e abórdase nun contexto xeral, dentro da teoría de integración estocástica en espazos de Banach UMD e do cálculo de Malliavin desenvolto nese marco. A continuación, analízase un estudo numérico aplicado a carteiras de pólizas de seguros a prazo, cando se considera un LIBOR Market Model para describir a evolución temporal das tasas de xuro forward.

Doutra banda, o estudo da estabilidade asintótica dunha clase dada de estimadores do risco, cando se amplía o conxunto de datos, proporciona un obxectivo adicional ao presente traballo. Neste sentido, unha noción refinada de robustez, que se aplica cando se manexan medidas de probabilidade aleatorias, caracterízase en termos das propiedades de converxencia verificadas polos procesos empíricos asociados a unha sucesión dada de variables aleatorias estacionarias.

CHAPTER 1

Introduction

Over the past decades, the life insurance industry has undergone major changes driven by the evolution of the economic environment and a combination of different advances both in technology and financial practices. Today's life insurance market is far more complex than in the past and it offers a wide range of different contracts, which are difficult to differentiate without an expert knowledge. As a result of these developments, assets and liabilities cash flow matching had has significantly increased and insurance sector has been made progressively more connected to financial markets. Nowadays, insurance companies deal with a large number of different securities, which are traded in order to hedge their exposure.

Even on a regulatory level, the life insurance sector experienced dramatic transformations aimed at promoting overall welfare. In order to oversee and enforce the regulation within the European Union insurance sector, the European Insurance and Occupational Pension Authority (EIOPA) was created in 2010 as a part of the European System of Financial Supervision. It constitutes an independent entity which provides advice to the European Commission, the European Parliament and the Council of the European Union. Some changes in the regulation within the life insurance sector have been also promoted by the European Court of Justice, with the major Test Achats case ruled on March 2011, which prohibits the gender-based discrimination in policy pricing.

Under the Solvency II Directive issued by the European Union in November 2009 and entered into force in January 2016, insurance companies are required to store a sufficient amount of funds aimed at reducing the social cost of potential losses and preventing future crisis. The required capital reflects the risk faced by the insurer and it is to be based on a market-consistent valuation approach of the assets and the liabilities in the balance sheet.

According to this framework, market risk management plays a central role within the insurance sector as well. In this respect, a major source of risk is provided by the fluctuation of the interest rate term structure, since most of the products traded by the insurers in order to hedge its exposure range typically from bonds or bond-type instruments to complex interest rates securities.

Many of the updates that has been undergone within the risk management practices give rise to a number of problems in the life insurance industry, which involve and promote the development of sophisticated models and mathematical tools.

1.1 LINES OF RESEARCH AND OBJECTIVES OF THE THESIS

This thesis is the result of the research activities conducted as a PhD student at *Universidade da Coruña* and Afi in Madrid, in the period between October 2015 and September 2018.

This work has been developed under the auspices of the *Wakeupcall* project funded in the EU Horizon2020 framework, which has been promoted in order to contribute in the advances that have been taken place within the financial practices as a consequences of the 2008 financial crisis.

Based on the objectives of the project, the following lines of research have been derived.

1. Robust risk estimation. As a consequence of the 2008 financial crisis, risk management experienced a number of transformations caused by the changes in regulation and the development of new financial practices. As a result, the modern theory of risk measures started with the seminal work by Artzner et al. [7] has been strongly pushed forward over the last decade.

According to the standard approach, the downside risk of a certain exposure is assessed on the basis of the past information. As a consequence, the asymptotic stability of the risk estimators when enlarging the historical dataset provides a fundamental issue.

Based on these features, the following objectives have been derived.

1.1 *Qualitative robustness.* Besides statistical consistency, Cont et al. [44] highlighted that the notion qualitative robustness introduced by Hampel [97, 98] and Huber [102] is a desirable property of asymptotic stability when dealing with risk estimators. On the other hand, an overall framework according to which statistical consistency and qualitative robustness may be regarded from a common perspective is missed in the existing literature.

    The main goal is to assess the results regarding the asymptotic stability of a given set of estimators in terms of stationary sequences of random variables. Further, the development of a refined version of qualitative robustness derived from the notion of statistical consistency and that applies when dealing with random probability measures constitutes a primary objective.

1.2 *Topological issues.* Cont et al. [44] and Kou et al. [113] highlighted that some commonly used risk functionals fail to be continuous with respect to the weak topology of measures. Briefly, the reason lies behind the fact that the weak topology is not sensitive enough to the tail behaviour of the distributions, which, by the way, is the main issue in risk analysis. Some authors [114, 115, 185, 186] suggested to overcome this lack by considering a proper refinement of the weak topology defined in terms of given functions that assess the displacement of the distributions on their tails.

Stating statistical consistency and qualitative robustness in terms of this topological refinement is a main task. On the other hand, the Borel σ-algebra generated by the weak topology of measures owns desirable properties that play a central role in this context, when dealing with many measurability issues. Thus, characterizing the Borel σ-algebra generated by the refined topology cited above represents a primary objective.

2. Portfolio representation. Recent updates in risk practices and regulations within the life insurance sector have enhanced the relevance of the stochastic models for the assets and liabilities management. In this respect, appropriate tools are needed in order to oversee the balance sheet of the company, which may involve portfolios of considerable size in the case of well-established life insurance firms.

In this dissertation we focus the attention on the liabilities of the company, which are mainly represented by the contracts sold to the clients. According to the new regulatory framework, insurance firms are allowed to reduce the scale of their portfolio of liabilities for analysis and management purposes, by considering only a certain group of products. Such a procedure is allowed only in the case when this scale reduction does not result in any misrepresentation of the risks underling the original portfolio.

Based on these features, the following objectives have been addressed.

2.1. *Portfolio representation.* The practice of the efficient portfolio reduction may be understood in connection with the problem of portfolio representation, in which one seeks to substitute a given portfolio by a second one that owns similar characteristics in terms of risks.

The primary objective in this context is to address this problem on a general level by minimizing a certain risk functional, which assesses the discrepancy between the two portfolios in terms of the stochastic variation of the underling risk factors over a given time horizon.

2.2. *Stochastic integration in Banach spaces.* The fluctuation of interest rates represents a main source of risk for any life insurance firm. According to the standard models, infinite dimensional dynamics are considered to capture the stochastic trend of the interest rate term structure over time. In this respect, many authors have developed a theory of bond portfolio in a infinite dimensional Hilbert space framework [39, 40, 69, 156].

In recent years, many results in stochastic calculus have been generalized to Banach spaces. In particular, Van Neerven, Veraar and Weis [174, 177, 178] recently advanced a theory for stochastic integration when dealing with the so-called UMD Banach spaces.

In this respect, developing a consistent portfolio theory in a Banach space framework represents a primary goal.

2.3. *Malliavin calculus.* Kettler, Prosk and Rubtsov [110] recently proposed a stochastic notion of bond duration by using Malliavin calculus for Gaussian random fields in a Hilbert space setup. The duration is a largely exploited tool when dealing in the fixed income market, since it provides the sensitivity of a given portfolio to the interest rates fluctuation.

Motivated by the developments in stochastic integration described above, some authors [127, 128, 149] extended a number of standard notions of Malliavin to a Banach spaces framework.

In this connection, assessing the problem of portfolio representation in terms of Malliavin calculus in Banach spaces represents a natural feature.

2.4. *Policy portfolio theory.* According to the previous arguments, addressing the problem of the policy portfolios scale reduction as a particular instance of the portfolio representation problem constitutes a major goal. For this purpose, the development of a consistent life insurance policy portfolio theory in Banach space is needed.

2.5. *Numerical issues.* The numerical minimization of the risk functional that we introduce in order to assess the portfolio representation problem may be computationally demanding, since it generally leads to a global optimization problem in high dimensions. In this respect, hybrid algorithms which combine a stochastic otpmization algorithm such as Simulated Anneaing [77] with a local gradient scheme [35, 123] may be used in order to address this problem. On the other hand, the numerical simulation of the interest rates dynamics may result in high computational costs, which may be addressed by considering efficient parallelization techniques.

   The development of efficient algorithms for the solution of the portfolio representation problem in life insurance and the analysis of their performance from the numerical point of view provide a main goal of this work. In this connection, the use of proper computer architectures represents a primary issue.

1.2 OUTLINE OF THE THESIS

This dissertation is organized in two different parts.

Part I refers to the studies that have been conducted within the modern theory of risk measures and robust risk estimation.

Chapter 2 briefly reviews the convergence results which motivate the research problems that are addressed in the following chapters. Particular emphasis is placed on the notion of weak topology and common forms of probabilistic symmetry, such as stationarity and exchangeability.

Chapter 3 presents a review of the modern theory of risk measures, based on a number of references that are cited therein. Particular attention is paid upon the notion of qualitative robustness in risk estimation.

The content of *Chapter 4* is mostly based on [75], except for the material in *Section 4.1*, which is not part of the cited work and which collects a personal development of different results, which have been found appealing for the sake of completeness. In this chapter a refined version of qualitative robustness is derived by the notion of statistical consistency in terms of random probability measures.

Part II deals with the problem of portfolio representation with particular emphasis on the life insurance sector.

Chapter 5 collects the main results within the theory of stochastic integration in a Banach space framework and the related literature. This review is based on a number of references cited therein. Although most of the results in this chapter are known, those that had not been found in a suitable form within the existing literature are discussed and proved in *Section 5.4*.

Chapter 6-7 contain original work and are based on [79]. In *Chapter 6* the problem of portfolio representation is addressed in terms of the theory of stochastic integration in Banach spaces while *Chapter 7* discusses this approach when dealing with life insurance portfolios. Special emphasis is placed on the case of portfolios composed by whole life insurance policies.

Chapter 8 is based on [76]. It consists in a numerical study of the approach discussed in the previous chapter applied to term insurance portfolios, by considering the LIBOR Market Model to describe the time evolution of the interest rate term structure.

Finally, a section discussing some open issues and ideas for further lines of research is presented.

Part I

Robust Risk Estimation

CHAPTER 2

Probabilistic Symmetries

Let $(E, \mathscr{E})$ be a Borel space. Hereafter, a *law* or *distribution* is a Borel probability measure defined on it. A random element in E is any E-valued random variable defined on a probability space $(\Omega, \mathscr{F}, \mathbb{P})$. A random sequence in E is a sequence of random elements in E. We shall write $\mathfrak{M}_1(E)$ to denote the entire family of laws on $(E, \mathscr{E})$. Moreover, for any measurable and real valued function f defined on $(E, \mathscr{E})$ and any law $\mu \in \mathfrak{M}_1(E)$ such that $\int_E |f| d\mu < \infty$, we shall write

$$\mu f \triangleq \int_E f d\mu.$$

Throughout this chapter, we shall characterize the topological structure of $\mathfrak{M}_1(E)$ and give an account of some convergence results that will play a relevant role later on. Moreover, we shall discuss the symmetries displayed by random sequences in E of major concern for the present work and that are to be regarded as a distributional forms of invariance that arise when considering certain families of measurable transformations.

2.1 WEAK TOPOLOGY AND EMPIRICAL PROCESSES

In this section, we assume E to be metrizable. Thus, we shall write d_E to denote the generic metric defined of it and consistent with its topological structure. In this case, we have that $\mathscr{E}$ coincides with the Bair σ-algebra generated by the the functionally open sets of the form $\mathscr{U}_f \triangleq \{x \in E : f(x) > 0\}$, where f is some continuous function on E, since in the present case E turns out to be *perfectly normal*, i.e. every closed set is the set of zeros of some continuous function on E (cf. [73], § 3.8). On the other hand, the σ-algebra $\mathscr{E}$ is also generated by the class of all bounded and continuous functions, and thus it consists in the smallest σ-algebra with respect to which any continuous function on E is measurable.

Recall that E is said to be a Polish space if there exists a metric d_E with respect to which it turns out to be complete and separable.

The weak topology. Let $\mathfrak{C}_b(E)$ denote the space of all real-valued continuous and bounded functions on E. Recall that it turns out to be a Banach space if endowed with the supremum norm $\|f\|_\infty \triangleq \sup_{x\in E} |f(x)|$, for any $f \in \mathfrak{C}_b(E)$.

Definition 2.1 (Weak Topology). The *weak topology* on $\mathfrak{M}_1(E)$ is the coarsest topology that renders continuous the map of the form $\mu \in \mathfrak{M}_1(E) \mapsto \mu f$, varying $f \in \mathfrak{C}_b(E)$.

As a result, any sequence $\mu_1, \mu_2, ...$ of laws in $\mathfrak{M}_1(E)$ converges to $\mu \in \mathfrak{M}_1(E)$ with respect to such topology, if and only if

$$\mu_n f \to \mu f, \qquad \text{for any } f \in \mathfrak{C}_b(E). \tag{2.1}$$

In this case, we say that the sequence $\mu_1, \mu_2, ...$ weakly converges to μ and we write $\mu_n \Rightarrow \mu$.

This notion of convergence can be restated in terms of the topological properties of the Borel sets in $\mathscr{E}$. At this end, for any Borel set $B \in \mathscr{E}$, $\bar{B}$ denotes the closure of B, int$B$ its interior (i.e. the largest open Borel set included in B) and $\partial B \triangleq \bar{B} \setminus \text{int}B$

the boundary of B. Recall that for any $B \in \mathscr{E}$ the boundary ∂B is closed. Moreover, we shall say that $B \in \mathscr{E}$ is a continuity set of μ if $\mu(\partial B) = 0$. It is worth to recall that the family of continuity sets forms an algebra included in $\mathscr{E}$ (cf. [21], Proposition 8.2.8).

The following result provides a well-known characterization for the weak convergence.

Theorem 2.1. *Let $(\mu_n)_n$ be a numerable family of laws in $\mathfrak{M}_1(E)$ and $\mu \in \mathfrak{M}_1(E)$, then the following statements are equivalent:*

i. $\mu_n \Rightarrow \mu$*;*

ii. For any open Borel set $U \in \mathscr{E}$, $\liminf_n \mu_n(U) \geq \mu(U)$*,*

iii. For any closed Borel set $C \in \mathscr{E}$, $\limsup \mu_n(C) \leq \mu(C)$*;*

iv. For any continuity set $B \in \mathscr{E}$ of μ, $\lim_n \mu_n(B) = \mu(B)$*.*

Proof. See, e.g., Theorem 6.1 in [144]. □

The following result gives a condition for which $\mathfrak{M}_1(E)$ endowed with the weak topology may be metrized as a separable metric space.

Theorem 2.2. *Endowed with the weak topology, the space $\mathfrak{M}_1(E)$ can be metrized as a separable space if and only if the space E is a separable metric space.*

Proof. See, e.g., Theorem 6.2 in [144]. □

Notice that the most natural form of smoothness is provided by the Lipschitz condition. In particular, we shall consider the family of bounded Lipschitz-continuous function $\mathfrak{BL}(E) \triangleq \{f \in \mathbb{R}^E : \|f\|_{\mathfrak{BL}} < \infty\}$, where $\|\cdot\|_{\mathfrak{BL}} \triangleq \|\cdot\|_\infty + \|\cdot\|_{\mathfrak{L}}$ has been defined by considering the norm $\|f\|_{\mathfrak{L}} \triangleq \sup_{u\neq v} |f(u) - f(v)|/d_E(u,v)$, for any Lipschitz-continuous function f on E. Clearly, $\|\cdot\|_{\mathfrak{BL}}$ turns out to be a norm on $\mathfrak{BL}(E)$. Moreover, one has that $fg \in \mathfrak{BL}(E)$ for any $f, g \in \mathfrak{BL}(E)$, and in particular $\|fg\|_{\mathfrak{BL}} \leq \|f\|_{\mathfrak{BL}}\|g\|_{\mathfrak{BL}}$, (cf. [63], Proposition 11.2.3).

Consider the application

$$\gamma(\mu,\nu) \triangleq \sup\{|(\mu-\nu)f| : \|f\|_{\mathfrak{BL}} \leq 1\}, \qquad \text{for any } \mu,\nu \in \mathfrak{M}_1(E), \tag{2.2}$$

where we let $(\mu-\nu)f \triangleq \mu f - \nu f$, that turns out to be a metric on $\mathfrak{M}_1(E)$, (cf. [63], Proposition 11.3.2.). Note that such a metric results to be similar to the distance generated by the *dual Lipschitz norm*

$$\|\mu\|_{\mathfrak{L}}^{\star} \triangleq \sup\Big\{|\mu f| : \|f\|_{\mathfrak{L}} \leq 1\Big\},$$

even if easily $\|\mu-\nu\|_{\mathfrak{L}}^{\star} \leq \gamma(\mu,\nu) \leq 2\|\mu-\nu\|_{\mathfrak{L}}^{\star}$, for any $\mu,\nu \in \mathfrak{M}_1(E)$. We shall refer to the metric (2.2) as the *Fortet-Mourier distance* on $\mathfrak{M}_1(E)$.

For any Borel set $B \in \mathscr{E}$, let $B^{\varepsilon} \triangleq \{x \in E : \inf_{y\in B} d_E(y,x) < \varepsilon\}$ denote the ε-hull of B. The application below

$$\pi(\mu,\nu) \triangleq \inf\{\varepsilon > 0 : \mu(B) \leq \nu(B^{\varepsilon}) + \varepsilon \text{ for any } B \in \mathscr{E}\}, \qquad \text{for any } \mu,\nu \in \mathfrak{M}_1(E), \tag{2.3}$$

provides another distance on the space $\mathfrak{M}_1(E)$, (cf. [63], Theorem 11.3.1). More precisely, we shall refer to (2.3) as the *Prohorov metric* on $\mathfrak{M}_1(E)$. Notice that, since $B^{\varepsilon} = (\bar{B})^{\varepsilon}$, we obtain an equivalent definition if B is required to be closed.

Remark. Observe that the Fortet-Mourier distance γ as well as the Prohorov distance π depend on the distance d_E generating the topological structure of E. Besides, if d'_E is a metric on E equivalent to d_E, then, with obvious notation, the corresponding distance π' is equivalent to π and the same for γ' and γ. This allows us to consider the consistent metric d that turns out to be more useful for our purposes. In particular, among all the distances consistent with the topology on E, there is one that is *totally bounded*, which will be convenient to use later on; see, e.g., Theorem 2.8.2 in [63]. On the other hand, note that the specific choice of the consistent metric d_E does not affect the separability of E.

The following result shows that, in presence of separability, the metrics (2.2) and (2.3) both metrize the weak topology on $\mathfrak{M}_1(E)$.

Theorem 2.3. *Assume E to be separable. Given a sequence $\mu_0, \mu_1, \dots$ of laws in $\mathfrak{M}_1(E)$, the following statements are equivalent*

i. $\mu_n \Rightarrow \mu_0$*;*

ii. $\mu_n f \to \mu_0 f$*, for any* $f \in \mathfrak{BL}(E)$*;*

iii. $\gamma(\mu_n, \mu_0) \to 0$*;*

iv. $\pi(\mu_n, \mu_0) \to 0$*.*

Proof. See, e.g., Theorem 11.3.3. in [63]. □

A family $\mathfrak{N} \subseteq \mathfrak{M}_1(E)$ is called *uniformly tight* if for any $\varepsilon > 0$ there exists a compact set $K_\varepsilon \subset E$ such that $\mu(K_\varepsilon) > 1 - \varepsilon$, for all $\mu \in \mathfrak{N}$. Moreover, a law $\mu \in \mathfrak{M}_1(E)$ is said *tight* if $\{\mu\}$ is uniformly tight.

We shall say that $\mathfrak{N}$ is *weakly relatively compact* if every sequence of laws in $\mathfrak{N}$ contains a weakly convergence subsequence, i.e. for any sequence $(\mu_n)_n$ of laws in $\mathfrak{N}$ there exist a subsequence $(\mu_{n_k})_k$ and a law $\mu \in \mathfrak{M}_1(E)$ (not necessarily in $\mathfrak{N}$) such that $\mu_{n_k} \Rightarrow \mu$. Moreover, $\mathfrak{N}$ is termed *weakly precompact* if its closure in the weak topology is compact.

Theorem 2.4. *Assume E to be a Polish space. If $\mathfrak{N}$ is a subset of $\mathfrak{M}_1(E)$, then the following statements are equivalent*

i. $\mathfrak{N}$ *is uniformly tight;*

ii. $\mathfrak{N}$ *is weakly relatively compact;*

iii. $\mathfrak{N}$ *is weakly precompact;*

iv. $\mathfrak{N}$ *is totally bounded for* π *and* γ*.*

Proof. See, e.g., Theorem 11.5.4. in [63]. □

Since every Cauchy sequence in a metric space is totally bounded, Theorem 2.4 directly leads to the following result.

Theorem 2.5 (Prohorov). *Assume E to be a Polish space. The space $\mathfrak{M}_1(E)$ is complete for π and γ.*

Convergence of empirical processes. We call *transition kernel* from Ω to E any map $\upsilon : \Omega \times \mathscr{E} \to \mathbb{R}^+$ such that

i the map $\upsilon(\cdot, B) : \omega \mapsto \upsilon(\omega, B)$ is $\mathscr{F}$-measurable for any $B \in \mathscr{E}$;

ii. the map $\upsilon(\omega, \cdot) : B \mapsto \upsilon(\omega, B)$ is a measure on $(E, \mathscr{E})$, for any $\omega \in \Omega$.

If λ is a finite measure on $(E, \mathscr{E})$ and Λ is a positive $\mathscr{F} \otimes \mathscr{E}$-measurable function on $\Omega \times E$, then the map

$$(\omega, B) \mapsto \int_B \Lambda(\omega, x)\lambda(dx), \qquad \text{for any } \omega \in \Omega \text{ and } B \in \mathscr{E},$$

defines a transition kernel from $(\Omega, \mathscr{F})$ to $(E, \mathscr{E})$. Moreover, if μ is finite on $(\Omega, \mathscr{F})$, every σ-bounded transition kernel υ from Ω to E, i.e. such that there exists a measurable partition $(E_n)_n$ of E for which $\omega \mapsto \upsilon(\omega, E_n)$ is bounded for any n, uniquely defines the measure α on the product space $(\Omega \times E, \mathscr{F} \otimes \mathscr{E})$ that is given below, (cf. [42], Theorem 6.11.),

$$\alpha(A \times B) \triangleq \int_A \upsilon(\omega\ , B)\mu(d\omega), \qquad \text{for } A \in \mathscr{F} \text{ and } B \in \mathscr{E}.$$

We shall say that a transition kernel υ is a *probability kernel*, if its range correspond to the unit interval on the real line, i.e. if $\upsilon(\omega, \cdot)$ is a probability measure on $(E, \mathscr{E})$, for any $\omega \in \Omega$. In this respect, we shall also refer to υ as a $\mathscr{F}$-measurable random measure on $(E, \mathscr{E})$.

Throughout this section, given a random sequence $\xi = (\xi_1, \xi_2, ...)$ in E we shall consider the family of probability kernels from Ω to E of the form

$$m_n(\omega, B) \triangleq \frac{1}{n} \sum_{i=1}^{n} \delta_{\xi_i(\omega)}(B), \qquad \text{for any } \omega \in \Omega \text{ and } B \in \mathscr{E}. \tag{2.4}$$

We usually refer to the sequence (2.4) as the empirical process directed by ξ.

Notice that if $f \in \mathfrak{C}_b(E)$, then $m_n f \to \mu f$ a.s., due to the strong law of large numbers. However the set of null measures on which the convergence fails depends on the specific function $f \in \mathfrak{C}_b(E)$. For this reason the following result is not trivial.

Theorem 2.6 (Varadarajan)**.** *Let E be separable and fix $\mu \in \mathfrak{M}_1(E)$. If the variables $(\xi_n)_n$ are independent with common law μ, then the empirical process (2.4) a.s. weakly converge to μ, i.e.*

$$\mathbb{P}\{\omega \in \Omega : m_n(\omega, \cdot) \Rightarrow \mu\} = 1.$$

Proof. See, e.g., Theorem 11.4.1 in [63]. □

Since almost everywhere pointwise convergence implies convergence in probability, under the same assumptions of the latter theorem, we get that

$$\chi(m_n, \mu) \triangleq \inf\{\varepsilon > 0 : \mathbb{P}(\pi(m_n, \mu) > \varepsilon) \leq \varepsilon\} \to 0, \qquad \text{as } n \to +\infty, \tag{2.5}$$

where χ denotes the Ky Fan metric associated to π, on the space of probability kernels from Ω to E, (cf. Theorem 9.2.2. in [63]).

Classically, the result discussed in Theorem 2.6 was initially introduced by Glivenko and Cantelli, when operating on the real line.

Theorem 2.7 (Glivenko-Cantelli)**.** *Let $\mu \in \mathfrak{M}_1(\mathbb{R})$ with cumulative function F. Then, $\|F_n(\omega, \cdot\,) - F\|_\infty \to 0$, as $n \to +\infty$, for $\mathbb{P}$-almost any $\omega \in \Omega$.*

Proof. See, e.g., Theorem 11.4.2. in [63]. □

Note that $m_n f \to \mu f$ a.s. as $n \to +\infty$, due to Theorem 2.6, when the variables $(\xi_n)_n$ that define the empirical process (2.4) are independent and with common law μ, for any $f \in \mathfrak{C}_b(E)$.

2.2 STATIONARITY AND EXCHANGEABLITY

Throughout this section, we discuss the main result that arise when considering certain forms of symmetry that characterize the distribution of the random sequences in E.

Stationarity. Fix a measure $\mu \in \mathfrak{M}_1(E)$ with respect to which the triple $(E, \mathscr{E}, \mu)$ turns out to be a σ-finite probability space, i.e. in such a way that there exists a countable measurable partition $(E_n)_n$ of E such that $\mu(E_n)$ is finite for any $n \geq 1$. A measurable transformation T from E into itself is called *measure preserving* if $\mu \circ T^{-1} \equiv \mu$. Thus, if ξ is a random element in E with distribution μ, the transformation T is *measure preserving* if and only if $\mathscr{L}(T\xi) = \mathscr{L}(\xi)$. Moreover, a random sequence $\xi \triangleq (\xi_n)_n$ in $(E, \mathscr{E})$ is said to be *stationary* if $\mathscr{L}(\Sigma\xi) = \mathscr{L}(\xi)$, where $\Sigma(x_n)_n \triangleq (x_{n+1})_n$ denotes the *shift* operator defined on the space $E^{\mathbb{N}}$ of the sequences in E. These notions appear to be strongly related, since T is measure preserving if and only if $(T^n\xi)_n$ is stationary, (cf. [107], Lemma 9.1).

Note that when dealing with the space $E^{\mathbb{N}}$, the measurable structure may be obtained by considering the product σ-algebra $\mathscr{E}^{\mathbb{N}}$, which is generated by the evaluation maps of the form $\pi_n : x \mapsto \pi_n x \triangleq x_n$, for any $x = (x_n)_n$ in $E^{\mathbb{N}}$.

Mutatis mutandis, the notion of stationarity extends naturally to random sequences indexed by $\mathbb{Z}$.

Throughout, let $\mathscr{I}_T$ be the entire family of T-inavtaint set, i.e. the measurable sets $I \in \mathscr{E}$ such that $T^{-1}(I) = I$. Note that, since the map T^{-1} preserves all the set operations, we get that $\mathscr{I}_T$ is a sub-σ-algebra of $\mathscr{E}$.

Theorem 2.8 (Birkhoff, von Neumann)**.** *Let T be a measurable transformation on E with associated T-invariant σ-algebra $\mathscr{E}_{inv(T)}$. Given a random element ξ in E, suppose that $(T^n\xi)_n$ is stationary. Then, for any $f \in L^p(E, \mathscr{E}, \mu)$, for some $p \geq 1$, we get*

$$m_n f \to \mathbb{E}[f(\xi)|\xi^{-1}\mathscr{I}_T], \qquad \text{a.s. and in } L^p, \text{ as } n \to +\infty, \tag{2.6}$$

where $(m_n)_n$ is the empirical process (2.4) obtained by setting $\xi_n \triangleq T^{n-1}\xi$, for $n \geq 1$, with the convention that $T^0 \triangleq \mathbb{I}_E$ represents the identity operator on E.

Proof. See, e.g., Theorem 9.6. in [107]. ☐

Notice that, under the hypotheses of Theorem 2.8, since $(T^n\xi)_n$ is stationary, and thus T preserves the measure μ, every random variable within the sequence $(\xi_n)_n$ admits the common distribution μ.

Ergodicity. A measure preserving transformation T from E into itself is said *ergodic* if the T-invariant σ-algebra $\mathscr{I}_T$ is μ-trivial, i.e. $\mu(B) \in \{0,1\}$ for any $B \in \mathscr{I}$. Depending on the specific viewpoint, we shall say that μ is ergodic with respect to T or that T is ergodic with respect to μ. In a natural way, a random element ξ in E is said *ergodic with respect to* T if its distribution on $(E, \mathscr{E})$ is ergodic for T. In such a case, the sub-σ-algebra $\xi^{-1}\mathscr{I}_T$ of $\mathscr{F}$ turns out to be $\mathbb{P}$-trivial.

On the other hand, a random sequence $(\xi_n)_n$ in E is said to be ergodic if its distribution on $(E^{\mathbb{N}}, \mathscr{E}^{\mathbb{N}})$ is ergodic with respect to the shift operator Σ, i.e. the shift-invariant σ-algebra $\mathscr{I}_\Sigma$ composed by the set $I \in \mathscr{E}^{\mathbb{N}}$ such that $\Sigma^{-1}(I) = I$ is trivial with respect to the distribution induced by $(\xi_n)_n$ on $(E^{\mathbb{N}}, \mathscr{E}^{\mathbb{N}})$. Moreover, the random elements in E that are ergodic with respect to some measure preserving transformation T are those and those only the random elements ξ such that the sequence $(T^n\xi)_n$ is ergodic with respect to Σ. In this case, every random sequence $(\Lambda \circ T^n\xi)_n$ is also ergodic for Σ, for any measurable operator Λ on E into itself (cf. [107], Lemma 9.5).

When ξ is assumed to be ergodic with respect to T, or equivalently when $(T^n\xi)_n$ is ergodic with respect to the shift operator Σ on $E^{\mathbb{N}}$, Theorem 2.8 boils down to

$$m_n f \to \mu f, \qquad \text{a.s. and in } L^p, \text{ as } n \to +\infty, \tag{2.7}$$

where μ denotes the distribution induced by ξ on $(E, \mathscr{E})$, since in such a case the sub σ-algebra $\xi^{-1}\mathscr{I}_\Sigma$ of $\mathscr{F}$ turns out to be $\mathbb{P}$-trivial.

Law of large numbers. Recall that the measure $\mathbb{P}$ is always trivial on the tail σ-algebra $\mathscr{T}_\xi \triangleq \cap_n \sigma\{\xi_k : k > n\}$, when $\xi_1, \xi_2, ...$ are independent random elements in $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$, due to the Kolmogorov's 0-1 law, (cf. [107], Theorem 2.13). As a result, since $\xi^{-1}\mathscr{I}_\Sigma \subset \mathscr{T}_\xi$, when assuming that $(\xi_n)_n$ are even i.i.d. and integrable and setting

$E = \mathbb{R}$, one has that Theorem 2.8 boils down to $n^{-1}(\xi_1 + ... + \xi_n) \to \mathbb{E}[\xi_1]$ a.s. and in L^1, as $n \to +\infty$. Hence, Theorem 2.8 can be regarded as a generalization of the Kolmogorov's law of large numbers.

Conditional probability. Given a random variable $\xi \in L^2(\Omega, \mathscr{F}, \mathbb{P})$ and a sub σ-algebra $\mathscr{C} \subseteq \mathscr{F}$, the conditional expectation $\mathbb{E}[\xi|\mathscr{C}]$ of ξ given $\mathscr{C}$ is defined as the Hilbert space orthogonal projection of ξ onto the linear subspace of $\mathscr{C}$-measurable random variables, i.e. $\mathbb{E}[(\xi - \mathbb{E}[\xi|\mathscr{C}])\eta] = 0$, for any bounded $\mathscr{C}$-measurable random variable η. This definition extends to any $\xi \in L^1(\Omega, \mathscr{F}, \mathbb{P})$ by continuity. It is worth to be highlighted that the conditional expectation is defined only up a null set, i.e. two version of the conditional expectation defined above coincide almost certainty with respect to the restriction $\mathbb{P}|_{\mathscr{C}}$ of $\mathbb{P}$ to $\mathscr{C}$.

A natural way to define a conditional probability given $\mathscr{C} \subseteq \mathscr{F}$ is obtained by setting $\mathbb{P}[A|\mathscr{C}] \triangleq \mathbb{E}[\mathbb{1}_A|\mathscr{C}]$, for any $A \in \mathscr{F}$, where this identity is to be understood in the a.s. sense with respect to $\mathbb{P}|_{\mathscr{C}}$. However, the $\mathbb{P}|_{\mathscr{C}}$-null sets on which the σ-additivity $\mathbb{P}[\cup_n A_n|\mathscr{C}] = \sum_n \mathbb{P}[A_n|\mathscr{C}]$ fails might depend on the sequence $A_1, A_2, ...$ in $\mathscr{F}$. As a result, the σ-additvity might fail in general, since the union of such sets may cover the entire space Ω. Nevertheless, this does not happen under some additional conditions, that we shall discus later.

Definition 2.2 (Conditional Probability). The map $\mathbb{P}[\,\cdot\,|\mathscr{C}](\cdot) : (\omega, A) \mapsto \mathbb{P}[A|\mathscr{C}](\omega)$, defined for any $\omega \in \Omega$ and $A \in \mathscr{F}$, is a *regular conditional probability* if the following assumptions are both true

i. For any $A \in \mathscr{F}$, the identity $\mathbb{P}[A|\mathscr{C}] \triangleq \mathbb{E}[\mathbb{1}_A|\mathscr{C}]$ holds a.s. with respect to $\mathbb{P}|_{\mathscr{C}}$, where the map $\omega \mapsto \mathbb{P}[A|\mathscr{C}](\omega)$ is $\mathscr{C}$-measurable;

ii. For $\mathbb{P}|_{\mathscr{C}}$-almost every $\omega \in \Omega$, $\mathbb{P}[\,\cdot\,|\mathscr{C}](\omega)$ is a probability measure on $(\Omega, \mathscr{F})$.

In the latter definition, the term "almost every" may be avoided by considering a $\mathbb{P}|_{\mathscr{C}}$-null set $C \in \mathscr{C}$, and thus by requiring that $\mathbb{P}[\,\cdot\,|\mathscr{C}](\omega)$ is a probability measure on $(\Omega, \mathscr{F})$ if $\omega \notin C$, and $\mathbb{P}[\,\cdot\,|\mathscr{C}](\omega) \triangleq \delta_{\bar{\omega}}(\cdot)$ when $\omega \in C$, for some $\bar{\omega} \in \Omega$ *a priori* fixed.

Conditional distribution and disintegration. Let now ξ be a random element in $(E, \mathscr{E})$. A regular conditional distribution of ξ given $\mathscr{C} \subset \mathscr{F}$ is defined as a version of the probability kernel $\mathbb{P}[\xi \in \cdot\, |\mathscr{C}]$ from $(\Omega, \mathscr{F})$ to $(E, \mathscr{E})$, i.e. a $\mathscr{C}$-measurable random probability measure on $(E, \mathscr{E})$.

More generally, we consider a measurable space $(T, \mathscr{T})$. If η is a random element in $(T, \mathscr{T})$, then the regular conditional distribution of ξ given η is defined as a random measure υ for which the equality $\upsilon(\eta, B) = \mathbb{P}[\xi \in B|\sigma(\eta)]$ holds a.s., for any $B \in \mathscr{E}$. Thus, υ turns out to be a probability kernel from $(T, \mathscr{T})$ to $(E, \mathscr{E})$. Throughout, we shall write $\mathbb{P}[\xi \in \cdot\, |\eta]$ is spite of $\mathbb{P}[\xi \in \cdot\, |\sigma(\eta)]$. Note that in the case when ξ is $\sigma(\eta)$-measurable or independent of η, one gets that $\mathbb{P}[\xi \in \cdot\, |\eta]$ reduces to $\mathbb{1}_{\{\xi \in \cdot\,\}}$ or $\mathbb{P}(\xi \in \cdot\,)$ respectively. The following result turns out to be crucial.

Theorem 2.9 (*Disintegration*)**.** *Fix a sub-σ-algebra $\mathscr{C}$ of $\mathscr{F}$ and consider two measurable spaces $(E, \mathscr{E})$ and $(T, \mathscr{T})$. Let ξ be a random element in $(E, \mathscr{E})$ and let υ be a regular version of $\mathbb{P}[\xi \in \cdot\, |\mathscr{C}]$. Furthermore, let η be a $\mathscr{C}$-measurable random element in $(T, \mathscr{T})$. Then*

$$\mathbb{E}[f(\xi, \eta)|\mathscr{C}] = \int_E f(x, \eta)\upsilon(dx), \qquad a.s. \tag{2.8}$$

for any measurable function f defined on $E \times T$ such that $\mathbb{E}|f(\xi, \eta)| < \infty$.

Proof. See, e.g., Theorem 5.4 in [107]. □

There exist many conditions that guarantee the existence of a regular version of a probability measure given a certain sub-σ-algebra. Since we deal with Borel metric space, we recall the following result.

Theorem 2.10. *Let $(E, \mathscr{E})$ a Borel space and μ a measure on it. Then, for any sub-σ-algebra $\mathscr{C}$ of $\mathscr{E}$, there always exists a regular version of the conditional law of μ given $\mathscr{C}$.*

Proof. See, e.g., Corollary 10.4.6 in [22]. □

Exchangeablility. Classically, a natural form of probabilistic symmetry is obtained by considering a law invariance under permutations. Let $\mathfrak{I}$ be a countable set of indices. A family $\{\xi_i : i \in \mathfrak{I}\}$ of random elements in $(E, \mathscr{E})$ is said to be *exchangeable* if

$$\mathscr{L}\{\xi_i : i \in \mathfrak{G}\} = \mathscr{L}\{\xi_{\pi_{\mathfrak{G}}(i)} : i \in \mathfrak{G}\}$$

for any finite family of indices $\mathfrak{G} \subset \mathfrak{I}$ and any permutation $\pi_{\mathfrak{G}}$ on it.

Note that every sequence of independent and identically distributed random variables is trivially exchangeable. Let $E^{\mathfrak{I}}$ be the family of the sequences in E indixed by $\mathfrak{I}$, and $\mathscr{E}^{\mathfrak{I}}$ the product σ-algebra generated by the projection maps $\pi_i : x \mapsto x_i$, for any $x \in E^{\mathfrak{I}}$, varying $i \in \mathfrak{I}$. For any random measure υ on $(E, \mathscr{E})$, we write $\upsilon(\omega, \cdot)^{\mathfrak{I}}$ to denote the product measure on $(E^{\mathfrak{I}}, \mathscr{E}^{\mathfrak{I}})$ defined by $\upsilon(\omega, \cdot)$ in the common way, for any fixed $\omega \in \Omega$. The following definition will play a relevant role later on.

A sequence $\{\xi_i : i \in \mathfrak{I}\}$ of random elements in $(E, \mathscr{E})$ is said *conditionally i.i.d.* if

$$\mathbb{P}[\xi \in \cdot \,|\mathscr{C}] = \upsilon^{\mathfrak{I}}, \qquad \text{a.s.} \tag{2.9}$$

for some sub-σ-algebra $\mathscr{C}$ of $\mathscr{F}$, and some random probability measure υ on $(E, \mathscr{E})$.

Note that, since υ turns out to be $\mathscr{C}$-measurable, the identity (2.9) is still valid when replacing $\mathscr{C}$ with $\sigma(\upsilon)$, i.e. the smallest σ-algebra on Ω with respect to with υ turns out to be a random measure on $(E, \mathscr{E})$.

Theorem 2.11 (de Finetti, Ryll-Nardzewski)**.** *Let ξ be a numerable random sequence in the Borel space $(E, \mathscr{E})$. Then, it is exchangeable if and only if $\mathbb{P}[\xi \in \cdot \,|\upsilon] = \upsilon^{\mathbb{N}}$ almost surely with respect to $\mathbb{P}$, for some random probability measure υ on $(E, \mathscr{E})$.*

Proof. See, e.g., Theorem 9.16 in [107]. □

Moreover, we refer to [5, 108, 111], for the proof of Theorem 2.11 by martingale arguments.

Hereafter, we shall refer to υ in Theorem 2.11 as the *directing random measure* associated to the sequence ξ. Moreover, for any numerable exchangeable random sequence, the associated directing random measure turns out to be a.s. unique (see, e.g., Proposition 1.4 in [108]).

0-1 laws. Recall that the shift-invariant σ-algebra $\mathscr{I}_\Sigma$ is defined as the collection of those measurable sets in $\mathscr{E}^\mathbb{N}$ that are invariant under the action of the shift map on $E^\mathbb{N}$. Recall also that it turns out to be μ-trivial for any distribution μ induced on $(E^\mathbb{N}, \mathscr{E}^\mathbb{N})$ by an ergodic sequence of random elements in $(E, \mathscr{E})$.

In a similar way, it is possible to introduce the measurable structure on $E^\mathbb{N}$ that naturally arises when dealing with permutations. In order to make this precise, note that any finite permutation π of $\mathbb{N}$, i.e. a map $\pi : \mathbb{N} \to \mathbb{N}$ such that $\pi(n) = n$ for all but finitely many $n \in \mathbb{N}$, induces a permutation map T_π on $E^\mathbb{N}$, by setting $T_\pi(x) \triangleq (x_{\pi(n)})_n$. IN this respect, a measurable set $I \in \mathscr{E}^\mathbb{N}$ is said *symmetric* if $T_\pi^{-1}(I) = I$, for any finite permutation π of $\mathbb{N}$. The family $\mathscr{P}$ of the symmetric sets is easily seen to form a sub-σ-algebra of $\mathscr{E}^\mathbb{N}$. In particular, we shall refer to $\mathscr{P}$ as the *permutation invariat* σ-algebra in $E^\mathbb{N}$.

The tail σ-algebra associated to the sequence ξ is defined as $\mathscr{T}_\xi \triangleq \cap_n \sigma\{\xi_k : k > n\}$. The following result shows that these σ-algebras are equivalent in the exchangeable setup.

Theorem 2.12 (Hewitt & Savage, Olshen)**.** *Let ξ be an exchangeable and numerable random sequence in a Borel space $(E, \mathscr{E})$ with directing random measure υ. Then,*

$$\sigma(\upsilon) = \xi^{-1}\mathscr{I}_\Sigma = \xi^{-1}\mathscr{P} = \mathscr{T}_\xi. \tag{2.10}$$

Moreover, all these σ-algebras are $\mathbb{P}$-trivial when dealing with a sequence ξ of i.i.d. random elements.

Proof. See, e.g., Corollary 1.6. in [108]. □

Strong Stationarity. Generally, stationarity and exchangeability are not equivalent since stationarity represents a stronger notion of distributional symmetry. Nevertheless, it is possible to restate the notion of exchangeability in terms of a stronger form of stationarity that looks at shift operators of random order.

Let $(\mathscr{F}_n)_n$ be a filtration in $\mathscr{F}$. We shall say that an $\mathscr{F}_n$-adapted random sequence ξ in $(E, \mathscr{E})$ is *$\mathscr{F}_n$-exchangeable* if every shifted sequence $\Sigma_n \xi \triangleq (\xi_n, \xi_{n+1}, ...)$ is conditionally exchangeable given $\mathscr{F}_n$, for any $n \geq 1$. In addition, we say that ξ is *strongly stationary*, or more properly *$\mathscr{F}_n$-stationary* if $\mathscr{L}(\Sigma_\tau \xi) = \mathscr{L}(\xi)$, for any finite stopping time τ associated to the filtration $(\mathscr{F}_n)_n$.

Theorem 2.13. *Let ξ be a numerable random sequence in a Borel space $(E, \mathscr{E})$ adapted to some filtration $(\mathscr{F}_n)_n$. Then, the following statements are equivalent*

i. ξ is $\mathscr{F}_n$-stationary;

ii. ξ is $\mathscr{F}_n$-exchangeable;

iii. The prediction sequence $\pi_n \triangleq \mathbb{P}[\Sigma_n \xi \in \cdot \, | \mathscr{F}_n]$, for $n \geq 0$, is a measure-valued martingales, i.e. $\big(\pi_n(B)\big)_n$ is a real-valued $\mathscr{F}_n$-martingales, for any $B \in \mathscr{E}$.

Proof. See, e.g., Proposition 9.18 in [107]. $\square$

CHAPTER 3

Risk Measures

In the wake of the financial crunch started in August 2007, the development of efficient models for the downside risk estimation has been has been increased of importance across the financial institutions. The problem naturally extended to the academic world, since it appears challenging also from the theoretical point of view.

In the new era of quantitative risk management, the downside risk is assessed as a function of the exposure worth at the end the trading period. These type of functions are nowadays called risk measures. According to the seminal work of Artzner et al. [7], any risk measure is defined by considering a certain class of properties that it should satisfy in order to be consistent from both the theoretical and the practical point of view. Such a set of properties forms the axiomatic system that defines the class of the so-called *coherent risk measures*.

This chapter presents a review of the modern theory of risk measures, with a main focus on the robust risk estimation.

3.1 AXIOMATIC AND SETUP

Let Ω be a set in which any element represents a certain scenario. A financial position is generally described in terms of some map $\xi : \Omega \to \mathbb{R}$, where $\xi(\omega)$ represents the discounted net worth of the position at the end of the trading period, if the scenario $\omega \in \Omega$ occurs. We follow the standard approach by assuming that positive values for ξ stand for gains while negative values describe the losses.

Monetary risk measures. Throughout, we assume a set $\mathfrak{X}$ of functions $\xi : \Omega \to \mathbb{R}$ to be fixed.

Definition 3.1 (Monetary Risk Measure). A functional $\rho : \mathfrak{X} \to \mathbb{R}$ is said a *monetary risk measure* if the following properties are satisfied,

A1. (Monotonicity) $\rho(\xi) \leq \rho(\eta)$, for any $\xi, \eta \in \mathfrak{X}$ such that $\xi(\omega) \geq \eta(\omega)$ for any $\omega \in \Omega$,

A2. (Cash Invariance) $\rho(\xi + m) = \rho(\xi) - m$, for any $\xi \in \mathfrak{X}$ and $m \in \mathbb{R}$.

The properties required in Definition 3.1 admit a natural interpretation. On one hand, A1 encodes the fact that the downside risk reduction is guaranteed when the financial profile is increased. On the other hand, A2 encodes the fact that the overall risk of the position is reduced by an amount m, when the same quantity m is added to the position and invested in a risk-free manner. From a regulatory point of view, such capital can be read as a security fund against the financial crunch.

Moreover, note that any monetary risk measure $\rho : \mathfrak{X} \to \mathbb{R}$ satisfies the following condition

$$\rho(\xi + \rho(\xi)) = 0, \qquad \text{for any } \xi \in \mathfrak{X}.$$

Throughout, we shall define the acceptance set associated to some risk measure $\rho : \mathfrak{X} \to \mathbb{R}$ as the set $\mathscr{A}_\rho \subseteq \mathfrak{X}$ defined by

$$\mathscr{A}_\rho \triangleq \{\xi \in \mathfrak{X} : \rho(\xi) \leq 0\}. \tag{3.1}$$

On the other hand, given any $\mathscr{A} \subset \mathfrak{X}$ consider the application

$$\rho_{\mathscr{A}}(\xi) \triangleq \inf\{m \in |m + \xi \in \mathscr{A}\}, \qquad \text{for any } \xi \in \mathfrak{X}. \tag{3.2}$$

Definition 3.2 (Convex Risk Measure). A monetary risk measure $\rho : \mathfrak{X} \to \mathbb{R}$ is usually called a *convex risk measure*, if it is convex, i.e.

A3. (Convexity) $\rho(\lambda\xi + (1-\lambda)\eta) \leq \lambda\rho(\xi) + (1-\lambda)\rho(\eta)$, for any $\xi, \eta \in \mathfrak{X}$ and any $\lambda \in (0,1)$.

Property A3 in Definition 3.2 stands for the fact that if one diversifies the investment, the risk of the overall position should be reduced.

Proposition 3.1. *Let $\rho : \mathfrak{X} \to \mathbb{R}$ be a convex risk measure, then $\rho \equiv \rho_{\mathscr{A}_\rho}$ and $\mathscr{A}_\rho$ is convex and non-empty.*

Proof. See, e.g., Proposition 4.6 in [83]. □

Proposition 3.1 says that a convex risk measure may be regarded as the minimal amount of capital to put aside in order to make the position safe.

Definition 3.3 (Coherent Risk Measure). A monetary risk measure is said a *coherent risk measure* if it satisfies,

A4. (Positive Homogeneity) $\rho(\lambda\xi) = \lambda\rho(\xi)$, for any $\xi \in \mathfrak{X}$ and $\lambda > 0$,

A5. (Subadditivity) $\rho(\xi + \eta) \leq \lambda\rho(\xi) + (1-\lambda)\rho(\eta)$, for any $\xi, \eta \in \mathfrak{X}$ and $\lambda \in (0,1)$.

Property A4 in Definition 3.3 encodes essentially the belief that the risk must increase in the dimension of the exposure. However, risk could grow non linearly as the size of position increases, when considering illiquid portfolio.

Law invariance. Throughout, we fix a complete probability space $(\Omega, \mathscr{F}, \mathbb{P})$. Thus, we assume $\mathfrak{X}$ to be a set of real-valued random variables defined on it.

Definition 3.4 (Law Invariance). A risk measure $\rho : \mathfrak{X} \to \mathbb{R}$ is said to be *law-invariant* if $\rho(\xi) = \rho(\eta)$ for any couple of random elements $\xi, \eta \in \mathfrak{X}$ such that $\mathscr{L}(\xi) = \mathscr{L}(\eta)$.

Let $\mathfrak{M}(\mathfrak{X}) \triangleq \{\mathscr{L}(\xi) : \xi \in \mathfrak{X}\}$. Any law-invariant risk measure $\rho : \mathfrak{X} \to \mathbb{R}$ naturally defines the functional $\mathfrak{R}_\rho : \mathfrak{M}(\mathfrak{X}) \to \mathbb{R}$ given by

$$\mathfrak{R}_\rho(\mu) \triangleq \rho(\xi), \qquad \text{for any } \xi \in \mathfrak{X} \text{ such that } \mathscr{L}(\xi) = \mu, \tag{3.3}$$

We shall refer to the functional defined by the identity (3.3) as the *distribution based* risk functional associated to the measure ρ.

Definition 3.5 (Stochastic Dominance). Given a couple of real valued random variable ξ_1 and ξ_2, we call stochastic dominance of the first (resp. second) kind the order relation denoted by $\xi_1 \succeq_{(1)} \xi_2$ (resp. $\xi \succeq_{(2)} \eta$) such that $\mu_1 f \geq \mu_2 f$, for any real valued and monotone function (resp. monotone and concave) $f : \mathbb{R} \to \mathbb{R}$, where $\mathscr{L}(\xi_i) = \mu_i$, for $i = 1, 2$.

Clearly, every monetary risk measure $\rho : \mathfrak{X} \to \mathbb{R}$ is law-invariant if and only if it respects the stochastic order of the first kind, i.e $\rho(\xi) \leq \rho(\eta)$ when $\xi \succeq_{(1)} \eta$. Moreover, the following result holds also true.

Theorem 3.1 (Song & Young). *Every law-invariant convex risk measure $\rho : \mathfrak{X} \to \mathbb{R}$ respects stochastic dominance of the both kind, that is $\rho(\xi) \leq \rho(\eta)$ when $\xi \succeq_{(i)} \eta$, for $i = 1, 2$. The converse is true when in addition ρ is comonotonic convex, i.e.*

$$\rho(\lambda\xi + (1-\lambda)\eta) \leq \lambda\rho(\xi) + (1-\lambda)\rho(\eta), \qquad \text{for any } \lambda \in (0, 1),$$

for any couple of comonotonic variables $\xi, \eta \in \mathfrak{M}(\mathfrak{X})$, i.e. such that $\xi = f(\zeta)$ and $\eta = g(\zeta)$ for some real valued non decreasing functions f and g defined on the real line, and some $\zeta \in \mathfrak{M}(\mathfrak{X})$.

Proof. See, e.g., Theorem 3.6 in [166]. □

As a result, when dealing with a law invariant risk measure ρ, properties A1 and A2 in Definition 3.1, may be restated in terms of the distribution-based functional $\mathfrak{R}_\rho$ by requiring that

A1. $\mathfrak{R}_\rho(\mu) \leq \mathfrak{R}_\rho(\nu)$ for any $\mu, \nu \in \mathfrak{M}(\mathfrak{X})$ such that $\mu \succeq_{(1)} \nu$;

A2. $\mathfrak{R}_\rho(\Sigma_m \mu) = \mathfrak{R}_\rho(\mu) - m$, for any $m \in \mathbb{R}$ and $\mu \in \mathfrak{M}(\mathfrak{X})$, where $(\Sigma_m \mu)(A) \triangleq \mu(A - m)$ denotes the shift operator on the real line.

3.2 CONVEX MEASURES OF RISK

Convex risk measures admit a robust dual representation that allows to regard the properties A1 and A2 in Definition 3.1 and the property A4 in Definition 3.3 from the analytical point of view.

Dual representation. Let $(\Omega, \mathscr{F})$ be a measurable space and $\mathfrak{X}$ a family of $\mathscr{F}$-measurable real valued functions defined on it. Further, we shall consider a family $\mathfrak{Y}$ of finite signed measures μ on $(\Omega, \mathscr{F})$ such that $\int_\Omega |\xi| d|\mu| < \infty$ for any $\xi \in \mathfrak{X}$, where we write $|\mu| \triangleq \mu^+ + \mu^-$ to denote the total variation measure defined via the Jordan decomposition $\mu = \mu^+ - \mu^-$.

We shall assume $\mathfrak{X}$ and $\mathfrak{Y}$ to be locally convex topological paired vector spaces by considering the duality pairing

$$\langle \mu, \xi \rangle \triangleq -\int_\Omega \xi(\omega)\mu(d\omega), \qquad \text{for any } \xi \in \mathfrak{X} \text{ and any } \mu \in \mathfrak{Y}. \tag{3.4}$$

In this connection, throughout we assume that for any $\xi \in \mathfrak{X} \setminus \{0\}$ there exist $\mu \in \mathfrak{Y}$ such that $\langle \mu, \xi \rangle \neq 0$ and conversely for every $\mu \in \mathfrak{Y} \setminus \{0\}$ there is some $\xi \in \mathfrak{X}$ such that $\langle \mu, \xi \rangle \neq 0$.

As a result, every continuous linear functional on $\mathfrak{X}$ admits the form $\xi \mapsto \langle \mu, \xi \rangle$, for some $\mu \in \mathfrak{Y}$, and similarly every continuous linear functional on $\mathfrak{Y}$ may be written as $\mu \mapsto \langle \mu, \xi \rangle$, for some $\xi \in \mathfrak{X}$. Thus, we are allowed to let $\mathfrak{Y} \triangleq \mathfrak{X}^\star$ be the topological dual of the space $\mathfrak{X}$. Besides, $\mathfrak{X}$ and $\mathfrak{X}^\star$ equipped with the relative strong topologies form paired spaces, when assuming $\mathfrak{X}$ to be a reflexive Banach space $\mathfrak{X}$. Recall that $\mathfrak{X}$ is reflexive if it is isometrically isomorphic to its second dual $\mathfrak{X}^{\star\star}$ by considering the evaluation map $J : \mathfrak{X} \to \mathfrak{X}^{\star\star}$ such that the linear functional $J\xi$ defined at $\xi \in \mathfrak{X}$ satisfies $J\xi(\mu) \triangleq \langle \mu, \xi \rangle$, for any $\mu \in \mathfrak{X}^\star$.

Let $\bar{\mathbb{R}} \triangleq \mathbb{R} \cup \{\pm\infty\}$. The following definition will play a relevant role later on.

Definition 3.6 (Proper Functional). A functional $\rho : \mathfrak{X} \to \bar{\mathbb{R}}$ is said *proper* if $\rho(\xi) > -\infty$ and its domain $\mathscr{D}(\rho) \triangleq \{\xi \in \mathfrak{X} : \rho(\xi) < +\infty\}$ is nonempty.

Classical duality theory provides the robust characterization of a generic functional in terms of the topological dual of its domain. Given a functional $\rho : \mathfrak{X} \to \bar{\mathbb{R}}$, the *conjugate* of ρ is defined as

$$\rho^\star(\mu) = \sup_{\xi \in \mathfrak{X}} \{\langle \mu, \xi \rangle - \rho(\xi)\}, \qquad \text{for any } \mu \in \mathfrak{Y}.$$

Similarly, the conjugate $\rho^{\star\star}$ of the second order is defined as

$$\rho^{\star\star}(\xi) = \sup_{\mu \in \mathfrak{Y}} \{\langle \mu, \xi \rangle - \rho^\star(\mu)\}, \qquad \text{for any } \xi \in \mathfrak{X}. \tag{3.5}$$

Theorem 3.2 (Fenchel-Moreau). *Suppose that the functional $\rho : \mathfrak{X} \to \bar{\mathbb{R}}$ is convex and proper, then $\rho^{\star\star} \equiv lsc(\rho)$, where $lsc(\rho)$ denotes the lower semicontinuous hull of ρ with respect to the topology of definition on $\mathfrak{X}$.*

Proof. See, e.g., Theorem 7.71 in [163]. □

As a direct consequence, if a convex and proper functional $\rho : \mathfrak{X} \to \bar{\mathbb{R}}$ is further assumed to be lower semicontinuous, then it admits the following representation

$$\rho(\xi) = \sup_{\mu \in \mathfrak{Y}} \{\langle \mu, \xi \rangle - \rho^\star(\mu)\}, \qquad \text{for any } \xi \in \mathfrak{X}. \tag{3.6}$$

Conversely, if the identity (3.6) holds for some function $\rho^\star$, then ρ is lower semicontinuous and convex. Moreover, since $\rho^\star$ is always proper when ρ is proper, lower semicontinuous and convex, one has that the identity (3.6) may be restated as follows

$$\rho(\xi) = \sup_{\mu \in \mathscr{D}(\rho^\star)} \{\langle \mu, \xi \rangle - \rho^\star(\mu)\}, \qquad \text{for any } \xi \in \mathfrak{X}, \tag{3.7}$$

where $\mathscr{D}(\rho^\star) \triangleq \{\mu \in \mathfrak{Y} : \rho^\star(\mu) < +\infty\}$ stands for the domain of the conjugate $\rho^\star$.

The following result characterizes the properties A1 and A2 in Definition 3.1 and the property A4 in Definition 3.3, when dealing with convexity.

Theorem 3.3. *Suppose that $\rho : \mathfrak{X} \to \bar{\mathbb{R}}$ is proper, lower semicontinuous and convex. Moreover, let $\mathscr{D}(\rho^\star) \triangleq \{\mu \in \mathfrak{Y} : \rho^\star(\mu) < +\infty\}$ be the domain of its conjugate. Thus, we have that*

i. Condition A1 in Definition 3.1 holds if and only if every $\mu \in \mathscr{D}(\rho^\star)$ is nonnegative;

ii. Condition A2 in Definition 3.1 holds if and only if $\mu(\Omega) = 1$, for any $\mu \in \mathscr{D}(\rho^\star)$;

iii. Condition A4 in Definition 3.3 holds if and only if the following representation

$$\rho(\xi) = \sup_{\mu \in \mathscr{A}} \langle \mu, \xi \rangle, \qquad \textit{for any } \xi \in \mathfrak{X}, \tag{3.8}$$

holds true, where $\mathscr{A} \triangleq \{\mu \in \mathfrak{Y} : \langle \mu, \xi \rangle < \rho(\xi), \text{for any } \xi \in \mathfrak{X}\}$.

Proof. See, e.g., Theorem 2.2. in [158]. □

Capital requirement and worst case scenario. The identity (3.8) provides a natural interpretation when dealing with coherent risk measures. In this respect, fix a probability measure $\mathbb{P}$ on $(\Omega, \mathscr{F})$ and set $\mathfrak{X} \triangleq L^p(\Omega, \mathscr{F}, \mathbb{P})$, for some $p \in (1, +\infty)$. Recall that, for $1 < p < +\infty$, the space $L^p(\Omega, \mathscr{F}, \mathbb{P})$ turns out to be reflexive and its topological dual $L^p(\Omega, \mathscr{F}, \mathbb{P})^\star$ is isometrically isomorphic to the space $L^{p^\star}(\Omega, \mathscr{F}, \mathbb{P})$, where $p^\star$ is the conjugate index of p satisfying $1/p + 1/p^\star = 1$.

As a result, the dual pairing (3.4) is regarded as

$$\langle \eta, \xi \rangle \triangleq -\int_\Omega \xi(\omega)\eta(\omega)\mathbb{P}(d\omega), \qquad \text{for any } \xi \in L^p(\Omega, \mathscr{F}, \mathbb{P}) \text{ and } \eta \in L^{p^\star}(\Omega, \mathscr{F}, \mathbb{P}),$$

by identifying any $\mu \in \mathfrak{Y}$ with a version of the Radon-Nikodyym derivative $d\mu/d\mathbb{P}$. Then, when dealing with a coherent risk measure $\rho : \mathfrak{X} \to \mathbb{R}$, the identity (3.8) is recast as follows

$$\rho(\xi) = \sup_{\eta \in \mathscr{A}} -\mathbb{E}[\xi\eta], \qquad \text{for any } \xi \in L^p(\Omega, \mathscr{F}, \mathbb{P}), \tag{3.9}$$

where $\mathscr{A} \triangleq \{\eta \in L^{p^\star}(\Omega, \mathscr{F}, \mathbb{P}) : -\mathbb{E}[\xi\eta] < \rho(\xi), \text{ for any } \xi \in L^p(\Omega, \mathscr{F}, \mathbb{P})\}$. Note that in the same way any function $\eta \in \mathscr{A}$ uniquely defines a probability measure μ on $(\Omega, \mathscr{F})$, such that

$$\int_\Omega \xi(\omega)\mu(d\omega) = \int_\Omega \xi(\omega)\eta(\omega)\mathbb{P}(d\omega), \qquad \text{for any } \xi \in L^p(\Omega, \mathscr{F}, \mathbb{P}). \tag{3.10}$$

Any variable $\eta \in \mathscr{A}$ may be regarded as the subjective assessment of the riskiness associated to a given exposure. Therefore, from the regulatory point of view, the capital requirement $\rho(\xi)$ as given by the identity (3.9) is computed as the expected loss in the worst case scenario, where the mean is computed with respect to any model μ whose Radon-Nikodym derivative belongs to the conjugate domain $\mathscr{A}$.

Spectral measures of risk. For any $\xi \in \mathfrak{X}$ and any $\alpha \in (0,1)$ we write $F_\xi^{-1}(\alpha) \triangleq \inf\{x \in \mathbb{R} : \mathbb{P}(\xi \leq x) \geq \alpha\}$ to denote the generalized inverse of the cumulative function associated to the law of ξ. Recall that, for any $\alpha \in (0,1)$, we call Value-at-Risk of level α the functional $\mathrm{V@R}_\alpha : \mathfrak{X} \to \mathbb{R}$ given by

$$\mathrm{V@R}_\alpha(\xi) \triangleq -F_\xi^{-1}(\alpha), \qquad \text{for any } \xi \in \mathfrak{X}. \tag{3.11}$$

On the other hand, we call Expected Shortfall of level α the functional $\mathrm{ES}_\alpha : \mathfrak{X} \to \mathbb{R}$ given by,

$$\mathrm{ES}_\alpha(\xi) \triangleq \frac{1}{\alpha}\int_0^\alpha \mathrm{V@R}_\beta(\xi)d\beta, \qquad \text{for any } \xi \in \mathfrak{X}. \tag{3.12}$$

Note that both the functionals $\mathrm{V@R}_\alpha$ and ES_α, for any $\alpha \in (0,1)$, are law invariant monetary risk measures. Moreover, Expected Shortfall is a coherent risk measure since it is a convex functional, and thus subadditive, cf. [3, 4].

Fix $m \in \mathbb{N}$ and note that, given $\alpha_i \in (0,1]$ and $\lambda_i \geq 0$ for any $i = 1, ..., m$ such that $\sum_{i=1}^m \lambda_i = 1$, the convex combination $\sum_{i=1}^m \lambda_i \mathrm{ES}_{\alpha_i}$ provides a law-invariant coherent risk measure. Further, given a finite set of law-invariant coherent risk measures $\rho_1, ..., \rho_n$, the measure $\rho \triangleq \max_{i=1,...,m} \rho_i$ is again a law-invariant coherent risk measure. These arguments are generalized by the following result.

Theorem 3.4 (Kusuoka)**.** *Let $(\Omega, \mathscr{F}, \mathbb{P})$ be a non atomic probability space and set $\mathfrak{X} = L^p(\Omega, \mathscr{F}, \mathbb{P})$, for some $p \geq 1$. Thus, let $\rho : \mathfrak{X} \to \mathbb{R}$ be a law invariant lower-semicontinuous coherent risk measure. Then, there exists a set $\mathfrak{N}$ of Borel probability measures on the unit interval of the real line such that*

$$\rho(\xi) = \sup_{\mu \in \mathfrak{N}} \int_0^1 \mathrm{ES}_\alpha(\xi)\mu(d\alpha), \qquad \textit{for any } \xi \in \mathfrak{X}.$$

Proof. See, e.g., Theorem 6.24 in [163]. □

For any Borel probability measure on the unit interval $[0, 1]$ of the real line define

$$\rho_\mu(\xi) \triangleq \int_0^1 \mathrm{ES}_\alpha(\xi)\mu(d\alpha), \qquad \text{for any } \xi \in \mathfrak{X}. \tag{3.13}$$

Any functional of the form (3.13) turns out to be additive on the class of comonotoic variables. i.e.

$$\rho_\mu(\xi + \eta) = \rho_\mu(\xi) + \rho_\mu(\eta) \tag{3.14}$$

for any couple of variables $\xi, \eta \in \mathfrak{X}$ such that $\xi \triangleq f(\zeta)$ and $\eta \triangleq g(\zeta)$, for some non-decreasing real valued functions f and g defined on the real line, and some $\zeta \in \mathfrak{X}$.

Comonotonicity is the strongest form of dependence that two random variables may display. It encodes the fact that no diversification benefits should be expected when two positions are totally positive dependent. Thus, while convexity formally encodes diversification, comonotonicity shows which are the limit cases when diversification may not lead to positive effects. The risk functionals of the form (3.13) have been introduced by Acerbi [1] and they are usually called spectral measures of risk.

The family of spectral measures of risk coincides with the convex hull generated by the one-parameter family $\{ES_\alpha : \alpha \in [0, 1]\}$, and it counts all those, and those only, the coherent law-invariant risk measures that turn out to be comonotonic, i.e. such that the relation (3.14) is always guaranteed for any couple of comonotonic variables.

3.3 ROBUST RISK ESTIMATION

Let $\mathfrak{X}$ be a class of real-valued random variables. Furthermore, fix a law invariant monetary risk measure $\rho : \mathfrak{X} \to \mathbb{R}$ and let $\mathfrak{R}_\rho$ to be the associated distribution based risk functional defined by (3.3).

According to the standard procedure, the downside risk of any exposure is assessed on the basis of the past information. More precisely, the distribution $\mu \in \mathfrak{M}(\mathfrak{X})$ that represents the discounted net worth of the position at the end of the trading period is calibrated via the historical data. Thus, the amount $\mathfrak{R}_\rho(\mu)$ represents an estimate of the downside risk related to the measure ρ.

Cont et. al [44] pointed out that risk can be properly assessed only when dealing with robust functionals, bay considering th Hampel's notion of qualitative robustness. It refers to the property of the risk estimate to be stable with respect to little changes affecting the distribution of the financial exposure. In this respect we refer to the seminal works of Hampel [97, 98] and Huber [102]. Other authors further developed the notion of qualitative robustness within the theory of risk estimation. We refer to the work of Krätschmer, Schied and Zähle [114, 115], in which a notion of qualitative robustness is introduced by considering sequence of independent and equally distributed random variables. Zähle [186, 184, 185] provided a robustness characterization for dependent random sequences.

Within the cited works, all the results are provided by considering a refinement of the weak topology of measure in terms of a certain gauge function that assess the displacement of the distributions of their tails. This overture is reasonable when dealing in risk estimation since the weak topology of measure lacks to be sensitive to the tails of the distributions. i.e. two laws might be quite close with respect to some metric relative to the weak topology and at the same time display a completely different behaviour on their tails. On the other hand, many common risk functionals fails to be qualitative robust with respect to the weak topology of measures.

Orlicz spaces and Orlicz hearts. We shall call *Young function* any left-continuous non decreasing convex function $\phi : [0, \infty) \to [0, \infty]$ such that $\phi(0) = 0$ and $\phi(x) \uparrow +\infty$ as $x \uparrow +\infty$. Note that thanks to the convexity, any Young function is continuous, except possibly at a single point, where it jumps to $+\infty$. We say that a Young function ϕ

satisfies the Δ_2-condition if

$$\phi(2x) \leq C\phi(x), \tag{3.15}$$

for some constant $C > 0$ and any large enough $x \in \mathbb{R}$.

We write $L^0(\Omega, \mathscr{F}, \mathbb{P})$ to denote the space of equivalence classes of real valued random variable defined on $(\Omega, \mathscr{F}, \mathbb{P})$, where two random variables are equivalent if they are equal a.s.

For any given Young function ϕ, the space $L^\phi(\Omega, \mathscr{F}, \mathbb{P}) \triangleq \{\xi \in L^0(\Omega, \mathscr{F}, \mathbb{P}) : \mathbb{E}[\phi(k|\xi|)] < \infty, \text{ for some } k > 0\}$ is called the *Orlicz Space* associated to ϕ. The Orlicz spaces represent a natural generalization to the Lebesgue setting. Moreover, for any Young function ϕ we have that $(L^\phi(\Omega, \mathscr{F}, \mathbb{P}), \|\cdot\|_\phi)$ is a Banach space when endowed with the Luxemburg norm defined as follows (See, e.g., Theorem 2.1.11 in [67]):

$$\|\xi\|_\phi := \inf\{\lambda > 0 : \mathbb{E}[\phi(|\xi|/\lambda)] \leq 1\}, \qquad \text{for any } \xi \in L^\phi(\Omega, \mathscr{F}, \mathbb{P}). \tag{3.16}$$

The space $H^\phi(\Omega, \mathscr{F}, \mathbb{P}) \triangleq \{\xi \in L^0(\Omega, \mathscr{F}, \mathbb{P}) : \mathbb{E}[\phi(k|\xi|)] < \infty, \text{ for any } k > 0\}$ is said the *Orlicz Heart* associated to ϕ. Here and in the sequel, we shall always assume any Young function ϕ to be finite, since $H^\phi(\Omega, \mathscr{F}, \mathbb{P}) = \{0\}$ and $L^\phi(\Omega, \mathscr{F}, \mathbb{P}) = L^\infty(\Omega, \mathscr{F}, \mathbb{P})$ whenever ϕ takes the value $+\infty$. It is easy to prove that

$$L^\infty(\Omega, \mathscr{F}, \mathbb{P}) \subset H^\phi(\Omega, \mathscr{F}, \mathbb{P}) \subset L^\phi(\Omega, \mathscr{F}, \mathbb{P}) \subset L^1(\Omega, \mathscr{F}, \mathbb{P}).$$

Moreover, these inclusions may be strict, since the following result holds true.

Theorem 3.5. *Let ϕ be a Young function. One has that $H^\phi(\Omega, \mathscr{F}, \mathbb{P})$ coincides with $L^\phi(\Omega, \mathscr{F}, \mathbb{P})$ if and only if the function ϕ satisfies the Δ_2-condition (3.15).*

Proof. See, e.g., Theorem 2.1.13 in [67]. □

The dual representation (3.9) provides the natural tool in order to study convex risk measure in terms of Orlicz hearts. Indeed, since $H^\phi(\Omega, \mathscr{F}, \mathbb{P}) \subset L^1(\Omega, \mathscr{F}, \mathbb{P})$ it is possible to study the properties of $\rho|_{H^\phi(\Omega, \mathscr{F}, \mathbb{P})}$ for the generic risk measure ρ defined on $L^1(\Omega, \mathscr{F}, \mathbb{P})$. Moreover, every finite valued convex functional defined on some Olicz heart $H^{\phi}(\Omega, \mathscr{F}, \mathbb{P})$ turns out to be continuous with respect to the Luxemburg norm (3.16). These arguments are collected in the following result.

Theorem 3.6 (Filipović & Svindald, Cheridito & Li)**.** *Let $\rho : L^{\infty}(\Omega, \mathscr{F}, \mathbb{P}) \to \bar{\mathbb{R}}$ be a low invariant and closed (i.e. proper and lower semicontinuous functional) convex risk measure, then*

i. For any $1 \leq p \leq +\infty$ there exists a unique low invariant closed convex functional $\bar{\rho}_p : L^p(\Omega, \mathscr{F}, \mathbb{P}) \to \mathbb{R} \cup \{+\infty\}$ such that $\bar{\rho}_1|_{L^{\infty}(\Omega, \mathscr{F}, \mathbb{P})} \equiv \rho$. Moreover,

$$\bar{\rho}_p(\xi) = \sup_{\eta \in \mathbf{L}^{p^\star}(\Omega, \mathscr{F}, \mathbb{P})} \{\langle \eta, \xi \rangle - \rho^\star(\eta), \} \qquad \text{for any } \xi \in L^p(\Omega, \mathscr{F}, \mathbb{P}),$$

where $\rho^\star$ is the conjugate of ρ. Furthermore, for any $1 \leq p \leq +\infty$ we get $\bar{\rho}_p \equiv \bar{\rho}_1|_{L^p(\Omega, \mathscr{F}, \mathbb{P})}$.

ii. If there exists a finite Young function ϕ such that $\bar{\rho}_1 : L^1(\Omega, \mathscr{F}, \mathbb{P}) \to \mathbb{R}$ is finite on $H^{\phi}(\Omega, \mathscr{F}, \mathbb{P})$, then $\bar{\rho}_1$ is continuous with respect to the Luxemburg norm $\|\cdot\|_{\phi}$.

Proof. Theorem 3.6 collects the results in [41] and [80]. □

Throughout, we fix a Young function ϕ and we set $\psi(x) \triangleq \phi(|x|)$, for any $x \in \mathbb{R}$. Notice that $\mathfrak{M}_1^{\psi}(\mathbb{R}) \triangleq \{\mu \in \mathfrak{M}_1(\mathbb{R}) : \mu\psi < \infty\}$ may be represented by the space $H^{\phi}(\Omega, \mathscr{F}, \mathbb{P})$, by identifying any couple of random variables with the same law. As a direct result, every law-invariant risk measure defined on $H^{\phi}(\Omega, \mathscr{F}, \mathbb{P})$ is completely described by the associated distribution-based risk functional $\mathfrak{R}_{\rho}$ defined on $\mathfrak{M}_1^{\psi}(\mathbb{R})$.

Given any sequence of real valued variables $\xi = (\xi_1, \xi_2, ...)$ defined on $(\Omega, \mathscr{F}, \mathbb{P})$, we recall that the sequence $m_1, m_2, ...$ of probability kernels from Ω to $\mathbb{R}$ defined by setting

$$m_n(\omega, B) \triangleq \frac{1}{n} \sum_{i=1}^{n} \delta_{\xi_i(\omega)}(B), \qquad \text{for any } \omega \in \Omega \text{ and } B \in \mathscr{B}(\mathbb{R}), \text{ varying } n \geq 1. \tag{3.17}$$

is called the empirical process directed by ξ.

Notice that, for any Young function ϕ, one has that $m_n(\omega, \cdot) \in \mathfrak{M}_1^{\psi}(\mathbb{R})$, for any $n \geq 1$ and any $\omega \in \Omega$. Thus, given any law invariant risk measure ρ defined on

$H^{\psi}(\Omega, \mathscr{F}, \mathbb{P})$, we write $\rho_1, \rho_2, ...$ to denote the sequence of real valued random variables defined by

$$\rho_n \triangleq \mathfrak{R}_{\rho}(m_n), \qquad \text{for any } n \geq 1. \tag{3.18}$$

We refer to (3.18) as the family of risk estimators associated to the measure ρ and directed by the sequence ξ. The following result gives sufficient condition for the convergence of the risk estimators (3.18) when dealing in the stationary and ergodic setup.

Theorem 3.7. *Let ϕ be a finite Young function and let $\xi_1, \xi_2, ...$ be an ergodic and stationary sequence in $H^{\phi}(\Omega, \mathscr{F}, \mathbb{P})$. If ρ is a law-invariant convex risk measure on $H^{\phi}(\Omega, \mathscr{F}, \mathbb{P})$, then $\rho_n \to \rho(\xi_1)$ a.s. as $n \to +\infty$.*

Proof. See Theorem 2.6. in [115]. □

Qualitative Robustness. Here and in the sequel, let $(\Omega, \mathscr{F}, \mathbb{P})$ be a probability space such that $\Omega \triangleq E^{\mathbb{N}}$ is the space of numerable sequences in E and $\mathscr{F} = \mathscr{E}^{\mathbb{N}}$ is the σ-algebra generated by the canonical projection $\xi_n : \omega \mapsto \xi_n(\omega) \triangleq \omega(n)$, for any $\omega \in \Omega$, varying $n \geq 1$. Throughout, we shall call *gauge function* on E any continuous function $\psi : E \to \mathbb{R}$ such that $\psi \geq 1$, everywhere on E. Moreover, given a gauge function ψ on E, we call ψ-weak topology the topology on $\mathfrak{M}_1^{\psi}(E)$ generated by the distance

$$d_{\psi}(\mu, \nu) \triangleq \pi(\mu, \nu) + \left| \int_E \psi d\mu - \int_E \psi d\nu \right|, \qquad \text{for any } \mu, \nu \in \mathfrak{M}_1^{\psi}(E).$$

Notice that the gauge function ψ assesses the displacement of the distribution in $\mathfrak{M}_1^{\psi}(E)$ on their tails. On the other hand, when ψ is bounded, one has that $\mathfrak{M}_1^{\psi}(E) = \mathfrak{M}_1(E)$, and the ψ-weak topology reduces to the standard weak topology on $\mathfrak{M}_1(E)$. We shall discuss the ψ-weak topology in Section 4.2.

Let M be a subspace of $\mathfrak{M}_1(E)$ such that all the empirical measures of the form $n^{-1} \sum_{i \leq n} \delta_{x_i}$, for $(x_1, ..., x_n) \in E^n$, endowed with a metric d_M. For any $\mu \in M$, we shall write $\mu^{\mathbb{N}}$ to denote the product measure on $(\Omega, \mathscr{F})$, i.e. such that the variables

$\xi_1, \xi_2, \ldots$ are independent and with the same common law μ. Moreover, when letting $(T, \mathscr{T})$ be a complete Borel space, for any $\tau : M \to T$ we shall write $\tau_n = \tau(m_n)$, for any $n \geq 1$.

Definition 3.7 (Qualitative Robustness). Fix $\mu \in M$. We say τ to be robust at μ with respect to (d_M, π) if for any $\varepsilon > 0$ there exist $\delta > 0$ and $n_0 \in \mathbb{N}$ such that

$$\text{if } \nu \in M \text{ with } d_M(\mu, \nu) < \delta, \quad \text{then} \quad \pi(\mu^{\mathbb{N}} \circ \tau_n^{-1}, \nu^{\mathbb{N}} \circ \tau_n^{-1}) < \varepsilon, \quad \text{for any } n \geq n_0. \tag{3.19}$$

A function $\tau : M \to T$ is said robust on M with respect to (d_M, π) if it is robust at every $\mu \in M$ with respect to (d_M, π).
When setting $M = \mathfrak{M}_1(E)$, $d_M = \pi$ the Prohorov measure on $\mathfrak{M}_1(E)$ given by (2.3), $T = \mathbb{R}^d$ and for any $\mu \in M$ the measure $\mathbb{P}_\mu = \mu^{\mathbb{N}}$ to be the prodcut measure on $(\Omega, \mathscr{F})$, Definition 3.7 boils down to the Hampel definition of robustness, as described in [98]. Cuevas [45] pushed forward the Hampel's notion of qualitative robustness, by letting T be a generic Polish space. Krätschmer, Shied and Zähle [114, 115] considered the notion of qualitative robustness by letting $M = \mathfrak{M}_1^{\psi}(\mathbb{R})$, for some gauge funtion ψ and thus letting $d_M = d_\psi$.

Note that, for any $\mu \in \mathfrak{M}_1(E)$, $m_n \Rightarrow \mu$ a.s. with respect to the measure $\mu^{\mathbb{N}}$, due to Theorem 2.6. Thus, since almost everywhere pointwise convergence implies convergence in measure, one has that

$$\kappa_\mu(m_n, \mu) \triangleq \inf \Big\{ \varepsilon > 0 : \mu^{\mathbb{N}} \Big\{ \omega \in \Omega : \pi\Big(m_n(\omega, \cdot\,), \mu\Big) > \varepsilon \Big\} \leq \varepsilon \Big\} \to 0, \qquad \text{as } n \to +\infty,$$

for any $\mu \in M$. These arguments lead to the following notion.

Definition 3.8 (UGC property). We shall say that $M \subseteq \mathfrak{M}_1(E)$ admits the *UGC property* with respect to d_M if

$$\sup_{\mu \in M} \inf \Big\{ \varepsilon > 0 : \mu^{\mathbb{N}} \Big\{ \omega \in \Omega : d_M\Big(m_n(\omega, \cdot\,), \mu\Big) > \varepsilon \Big\} \leq \varepsilon \Big\} \to 0, \qquad \text{as } n \to +\infty, \tag{3.20}$$

Definition 3.8 is taken from [114, 115], and the acronym *UGC* stands for *"Uniformly Glivenko Cantelli"*. However, it s worth to be highlighted that while in

Glivenko-Cantelli theorem the variable $\xi_1, \xi_2, \dots$ that drive the empirical process (3.17) are assumed to be independent and with the same common distribution, in Definition 3.8 we are not assuming independence.

The space $\mathfrak{M}_1(E)$ admits the UGC property with respect to π, (see, e.g., Lemma 2.4 in [136]). Thus, the space $\mathfrak{M}_1(E)$ admits the UGC property with respect to d_ψ if and only if

$$\sup_{\mu \in M} \inf \left\{ \varepsilon > 0 : \mu^{\mathbb{N}} \left\{ \omega \in \Omega : \left| \int_E \psi(x) dm_n(\omega, dx) - \int_E \psi(x) \mu(dx) \right| > \varepsilon \right\} \leq \varepsilon \right\} \to 0, \quad \text{as } n \to +\infty. \tag{3.21}$$

The following result provides a version of Hampel's theorem assessed in terms of the ψ-weak topology of measures.

Theorem 3.8. *Fix $M \subseteq \mathfrak{M}_1^\psi(\mathbb{R})$ and assume $\tau : M \to T$ to be continuous at $\mu \in M$ with respect to the ψ-weak topology. If M satisfies the UGC property then τ is robust at μ with respect to (d_ψ, π).*

Proof. See, e.g., Theorem 3.2 in [115]. □

Let d_T be some metric on T consistent with $\mathscr{T}$. For a given $\mu \in \mathfrak{M}_1(\mathbb{R})$, we say that the sequence $\tau_1, \tau_2, \dots$ is *weakly $\mathbb{P}_\mu$-consistent* with respect to d_T if $\inf\{\varepsilon > 0 : \mathbb{P}_\mu(d_T(\tau_n, \tau(\mu)) > \varepsilon) \leq \varepsilon\} \to 0$, as $n \to +\infty$. The following result provides a converse of Theorem 3.8.

Corollary 3.1. *Let $M \subseteq \mathfrak{M}_1^\psi(\mathbb{R})$ and fix a functional $\tau : E \to T$. Fix $\delta > 0$ and assume $\tau_1, \tau_2, \dots$ to be marginal robust at any $\nu \in M$ such that $d_\psi(\mu, \nu) < \delta$. If τ is robust at $\mu \in M$, then τ is continuous at μ with respect to the ψ-weak topology.*

Robust risk measures. Throughout, we fix a finite Young function ϕ and we set $\psi(x) = \phi(|x|)$, for any $x \in \mathbb{R}$. Notice that the function ψ provides a gauge function on E. The following result provides a useful characterization of the continuity of the risk functional $\mathfrak{R}_\rho$ for a given law invariant risk measure ρ, with respect to the ψ weak topology on $\mathfrak{M}_1^\psi(\mathbb{R})$.

Theorem 3.9 (Krätschmer, Schied & Zähle)**.** *Let ρ be a convex risk measure defined on $H^\phi(\Omega, \mathscr{F}, \mathbb{P})$. The functional $\mathfrak{R}_\rho : \mathfrak{M}_1^\psi(\mathbb{R}) \to \mathbb{R}$ is continuous with respect to the ψ-weak topology if and only if ϕ satisfies the Δ_2-condition.*

Proof. See, e.g., Theorem 2.8 in [115]. □

Similarly, a sequence $\mu_1, \mu_2, \ldots \in \mathfrak{M}_1^\psi(E)$ converges to a law μ in the ψ-weak topology if and only if $\|\xi_n - \xi\|_\phi \to 0$, where ϕ satisfies the Δ_2-condition and the variables $\xi, \xi_1, \xi_2, \ldots$ belong to $H^\psi(\Omega, \mathscr{F}, \mathbb{P})$ and are such that $\mathscr{L}(\xi_n) = \mu_n$ and $\mathscr{L}(\xi) = \mu$. In this respect, notice that these variables satisfying the previous condition always exists since $(\Omega, \mathscr{F}, \mathbb{P})$ is atomless, (cf. [115], Theorem 3.5).

Given a law-invariant convex risk measure $\rho : H^\phi(\Omega, \mathscr{F}, \mathbb{P}) \to \mathbb{R}$, let $\rho_1, \rho_2, \ldots$ be the associated sequence of estimators given by (3.18). Thus, Theorem 3.8 combined with Theorem 3.9 gives the following result.

Corollary 3.2. *Let $\rho : H^\phi(\Omega, \mathscr{F}, \mathbb{P}) \to \mathbb{R}$ be a law-invariant convex risk measure and fix $M \subseteq \mathfrak{M}_1^\psi(\mathbb{R})$ that satisfies the UGC property. The functional $\mathfrak{R}_\rho$ is robust on M with respect to (d_ψ, π) if and only if ϕ satisfies the Δ_2-condition.*

When dealing with financial risk management, for a given $M \subseteq \mathfrak{M}_1^\psi(\mathbb{R})$ we may regard any law in M as the expected distribution of a certain financial exposure in the future. In this respect, it is worth to be highlighted that continuity represents a property that the risk functional $\mathfrak{R}_\rho$ should reasonably display. Indeed, in this case, little forecasting mistakes can be controlled by injecting just contained amount of capital as safety fund. On the other hand, in this case the risk functional $\mathfrak{R}_\rho$ appears to be robust on any $M \subseteq \mathfrak{M}_1^\psi(\mathbb{R})$ that satisfy the UGC property with respect to (d_ψ, π). This means that little forecasting mistakes only result in little perturbations of the law of the estimators, provided that the amount of the processed date turns out to be large enough.

Recall that every convex law-invariant risk measure $\rho : L^\infty(\Omega, \mathscr{F}, \mathbb{P}) \to \mathbb{R}$ admits a unique extension $\bar{\rho}_1 : L^1(\Omega, \mathscr{F}, \mathbb{P}) \to \mathbb{R}$, and in the particular case in which $\bar{\rho}_1$ takes finite values on $H^\phi(\Omega, \mathscr{F}, \mathbb{P})$, for some finite Young function ϕ, one has that $\bar{\rho}_1$ turns out to be continuous with respect to the Luxemburg norm (3.16), thanks to Theorem 3.6. The following result extends these arguments.

Proposition 3.2. *Let $\rho : L^\infty(\Omega, \mathscr{F}, \mathbb{P}) \to \mathbb{R}$ be a law-invariant convex risk measure and hence let $\bar{\rho}_1 : L^1(\Omega, \mathscr{F}, \mathbb{P}) \to \mathbb{R}$ be its unique extension. If ϕ satisfies the Δ_2-condition, the following statements are equivalent.*

i. $\mathfrak{R}_{\bar{\rho}_1}$ is marginal robust on $\mathfrak{M}_1^\psi(\mathbb{R})$;

ii. $\mathfrak{R}_\rho$ is marginal robust on $\mathfrak{M}(L^\infty(\Omega, \mathscr{F}, \mathbb{P}))$;

iii. $\bar{\rho}_1$ takes finite values on $H^\phi(\Omega, \mathscr{F}, \mathbb{P})$.

Proof. See, e.g., Theorem 2.16 in [115]. □

Degrees of robustness. Here and in the sequel, for any $p \in [1, +\infty)$ we fix $\phi_p(x) \triangleq x^p/p$, for any $x \in \mathbb{R}^+$. Notice that ϕ_p is a finite Young function and that $H^{\phi_p}(\Omega, \mathscr{F}, \mathbb{P}) = L^p(\Omega, \mathscr{F}, \mathbb{P})$. When setting $\psi_p(x) = \phi_p(|x|)$, for any $x \in \mathbb{R}$, since ϕ_p satisfies the Δ_2-condition for any $p \in [1, +\infty)$, we have that for any convex law-invariant risk measure $\rho : L^p(\Omega, \mathscr{F}, \mathbb{P}) \to \mathbb{R}$, the functional $\mathfrak{R}_\rho : \mathfrak{M}_1^{\psi_p}(\mathbb{R}) \to \mathbb{R}$ is continuous with respect to the ψ_p-weak topology. In this setup, the ψ_p-weak topology is generated by the distance d_p given by the following identity, (see e.g. Theorem 7.12 in [180]),

$$d_p(\mu, \nu) \triangleq \inf \left\{ \left(\int |x_1 - x_2|^p \lambda(dx_1, dx_2) \right)^{1/p} : \lambda \in \mathfrak{M}_1(\mathbb{R} \times \mathbb{R}) \text{ with marginals } \mu \text{ and } \nu \right\}, \tag{3.22}$$

The metric d_p is usually called the *Monge-Wasserstein metric* of order p. The following result states that for $p = 1$, the Wasserstein metric (3.22) may be defined in terms of the norm (2.3).

Theorem 3.10 (Kantorovich-Rubinstein)**.** *For any $\mu, \nu \in \mathfrak{M}_1^{\psi_1}(\mathbb{R})$, one has $d_1(\mu, \nu) = \|\mu - \nu\|_{\mathfrak{L}}^\star$.*

Proof. See, e.g., Theorem 11.8.2 in [63]. □

For $p = 1$, the distance d_p is sometimes called the *Gini index*. It was studied by Gini around 1914, and later reviewed by Rachev [150]. This metric arises in the context of the so-called transportation problems initially posed by Monge [137].

The following definition is taken from [114].

DEFINITION 3.9 (Index of Qualitative Robustness). Given a law-invariant convex risk measure $\rho : L^\infty(\Omega, \mathscr{F}, \mathbb{P}) \to \mathbb{R}$, the quantity $\mathrm{iqr}(\rho)$ given by

$$\mathrm{iqr}(\rho) \triangleq \Big[\inf\{p \in (0, +\infty) : \mathfrak{R}_\rho \text{ is robust on } \mathfrak{M}\Big(L^\infty(\Omega, \mathscr{F}, \mathbb{P})\Big) \text{ with respect to } (d_p, \pi)\}\Big]^{-1} \tag{3.23}$$

is said the *index of qualitative robustness* associated to ρ.

For any given law-invariant convex risk functional $\rho : L^\infty(\Omega, \mathscr{F}, \mathbb{P}) \to \mathbb{R}$, the index (3.23) in Definition 3.9 may be regarded as a measure of the degrees of robustness displayed by $\mathfrak{R}_\rho$. Higher degrees of such an index should represent higher degrees of robustness. As a result, the dichotomy between classes of robust and not robust functionals may be replaced by a continuum of different degrees of robustness.

CHAPTER 4

Empirical Processes and Asymptotic Stability

In this chapter, we fix a Polish space E and we use $\mathscr{E}$ to denote its Borel σ-algebra. Moreover, we denote the space of probability measures on $(E, \mathscr{E})$ with $\mathfrak{M}_1(E)$. Let $(\Omega, \mathscr{F}, \mathbb{P})$ be a probability space in which we define $\Omega \triangleq E^{\mathbb{N}}$ to be the family of numerable sequences in E, and thus $\mathscr{F} \triangleq \mathscr{E}^{\mathbb{N}}$ to be the product σ-algebra on it, that coincides with the σ-algebra generated by the canonical projections $\xi_n : \omega \mapsto \xi_n(\omega) \triangleq \omega_n$, for any $\omega \in \Omega$, varying $n \geq 1$. We shall refer to the random sequence $\xi \triangleq (\xi_1, \xi_2, ...)$ as the *data process* and we shall call *empirical process* the family of random measures $(m_n)_n$ on $(E, \mathscr{E})$ defined by setting

$$m_n(\omega, B) \triangleq \frac{1}{n} \sum_{i=1}^{n} \delta_{\xi_i(\omega)}(B), \qquad \text{for any } \omega \in \Omega \text{ and } B \in \mathscr{E}. \tag{4.1}$$

Let Σ be the shift operator on E and hence define $\mathscr{I}_\Sigma$ to be the Σ-invariant σ-algebra of sets in $\mathscr{F}$, i.e. the class of $A \in \mathscr{F}$ such that $\Sigma^{-1}(A) = A$. Here and in the sequel, we write υ to denote a regular version of the conditional distribution $\mathbb{P}[\xi_1 \in \cdot | \xi^{-1} \mathscr{I}_\Sigma]$. It is worth to be highlighted that since E is Polish and $\mathscr{E}$ is its Borel

σ-algebra, and hence countably generated, the regular version υ always exists and it is essentially unique (see, e.g., Theorem 10.4.3 and 10.4.6 in [21]). According to this framework, we get that Theorem 2.8 combined with Theorem 2.9 leads to

$$m_n f \to \upsilon f, \qquad \text{a.s.}, \tag{4.2}$$

for any fixed $f \in \mathfrak{C}_b(E)$. Hence, when running f over the space $\mathfrak{C}_b(E)$, one should expect that $m_n \Rightarrow \upsilon$ a.s. However, this convergence result is not trivial, since the null set for the a.s. converge (4.2) may depend on the specific function f and the union of these sets may cover the entire space Ω.

In this chapter, the problem of the convergence of the sequence (4.1) is assessed via separability arguments by considering the weak topology of measures at first and then a certain refinement of it. These arguments allow to introduce a refined notion of robustness that applies when dealing with stationary sequences in terms of the random measure υ.

This chapter is based on the original work [75].

4.1 CONVERGENCE OF EMPIRICAL PROCESSES

In this section we study the convergence of the process (4.1) with respect to the weak topology on $\mathfrak{M}_1(E)$, which has been defined in §2.1 by considering the family of continuous and bounded functions on E. This characterization may be properly generalized by considering other functional spaces. In this respect, a generalized notion of weak topology may be introduced by considering classical duality arguments. However, we show that any family of bounded and uniformly continuous functions with some modulus of continuity actually generates a topology on $\mathfrak{M}_1(E)$ that coincides with the weak topology.

Glivenko-Cantelli classes. Recall that the space $\mathfrak{C}_b(E)$ endowed with the supremum norm $\|\cdot\|_\infty$ is, in general, not separable. Likewise, the space $\mathfrak{BL}(E)$ is, in general, not separable when considering the topology induced by the norm $\|\cdot\|_{\mathfrak{BL}(E)}$. Let us define $\mathfrak{BL}_1(E)$ to be the unit ball in $\mathfrak{BL}(E)$, i.e. the set of functions $f \in \mathfrak{BL}(E)$ such that $\|f\|_{\mathfrak{BL}(E)} \leq 1$.

Proposition 4.1. *If the distance d consistent with the topology on E is totally bounded, then the unit ball $\mathfrak{BL}_1(E)$ is separable for the supremum norm $\|\cdot\|_\infty$.*

Notice that the ball $\mathfrak{BL}_1(E)$ of $\mathfrak{BL}(E)$ depends on the actual distance d on E used in the definition of the norm $\|\cdot\|_{\mathfrak{BL}(E)}$. In the special case when d is indeed complete, then E appears to be compact, (cf. [63], Theorem 2.3.1). Thus, since the family $\mathfrak{BL}_1(E)$ is equicontinuous and uniformly bounded, Ascoli-Arzelá theorem (cf. [63], Theorem 2.4.7) ensures that $\mathfrak{BL}_1(E)$ is compact with respect to the topology induced by $\|\cdot\|_\infty$, and, consequently, separable. Besides, notice that the distance d in Proposition 4.1 might fail to be complete, since completeness is not a topological invariant.

The following proof of Proposition 4.1 is modelled on the ideas contained in the proof of Thereom 11.4.1 in [63].

Proof of Proposition 4.1. Let $\bar{E}$ be the completion of E with respect to the metric d, and define $\mathfrak{BL}_1(\bar{E})$ to be the unit ball in $\mathfrak{BL}(\bar{E})$ defined by the norm $\|\cdot\|_{\mathfrak{BL}(\bar{E})}$. Fix $g \in \mathfrak{BL}_1(E)$ and denote by $\bar{g}$ the unique extension of g (see, e.g. Proposition 11.2.3 in [63]) defined on $\bar{E}$ such that $\|\bar{g}\|_{\mathfrak{BL}(\bar{E})} = \|g\|_{\mathfrak{BL}(E)}$ and hence so that $\bar{g} \in \mathfrak{BL}_1(\bar{E})$.

Note that $\bar{E}$ is compact, since E is assumed to be totally bounded, (cf. [63], Theorem 2.3.1). Thus, there exists a subset $\mathfrak{N}$ of $\mathfrak{BL}_1(\bar{E})$ which is countable and dense in $\mathfrak{BL}_1(\bar{E})$ with respect to the supremum norm $\|\cdot\|_\infty$. As a direct result, given $g \in \mathfrak{BL}_1(E)$, for any $\varepsilon > 0$ one can find $f \in \mathfrak{N}$ so that $\|\bar{g} - f\|_\infty \leq \varepsilon$.

On the other hand, letting $f|_E$ be the restriction of f to the domain E, one has $\|g - f|_E\|_\infty \leq \|\bar{g} - f\|_\infty \leq \varepsilon$. Hence, the family $\mathfrak{N}|_E$ of the functions in $\mathfrak{N}$ restricted to E provides a dense and countable subset of $\mathfrak{BL}_1(E)$ in the norm $\|\cdot\|_\infty$. $\square$

Let μ be a distribution on $(E, \mathscr{E})$. A family $\mathfrak{F}$ of real-valued and $\mathscr{E}$-measurable functions defined on E is said to be a μ-*Glivenko-Cantelli* class (cf. [65, 169]) if

$$\sup\{|(m_n - \mu)f|,\ f \in \mathfrak{F}\} \to 0, \qquad \text{a.s. as } n \to +\infty, \tag{4.3}$$

when the variables $\xi_1, \xi_1, \ldots$ are independent with common distribution μ. Moreover, $\mathfrak{F}$ is termed a *universal Glivenko-Cantelly calss* if it is a μ-Glivenko-Cantelli class for any $\mu \in \mathfrak{M}_1(E)$. It is worth to be noted that when the underlying triple $(\Omega, \mathscr{F}, \mathbb{P})$ is assumed to be non-atomic, given any $\mu \in \mathfrak{M}_1(E)$, we can always find a sequence $\xi = (\xi_1, \xi_2, \ldots)$ of independent random elements in $(E, \mathscr{E})$ such that $\mathscr{L}(\xi_n) = \mu$, for any $n \geq 1$.

Notice that, according to Theorems 2.1 and 2.6 the unit ball $\mathfrak{BL}_1(E)$ of the space $\mathfrak{BL}(E)$ of bounded and Lipschitz continuous functions appears to be a universal Glivenko-Cantelli class, since (4.3) boils down to the convergence of the process $m_1, m_2, \ldots$ with respect to the distance (2.2).

Proposition 4.2. *If ξ is stationary, then $m_n \Rightarrow \upsilon$ a.s. as $n \to +\infty$.*

Recall that the metric γ defined by (2.2) actually depends on the specific metric d we consider on the space E. On the other hand, whenever two metrics d and d' generate the same topological structure on E, with obvious notation the metrics γ and γ' induces the same topology on $\mathfrak{M}_1(E)$. Thus, this allows us to define the norm $\|\cdot\|_{\mathfrak{BL}_1(E)}$ by considering the metric on E consistent with its topological structure, that turns out to be more convenient for our purposes. In particular, recall that there always exists a metric that renders E totally bounded, (cf. [63], Theorem 2.8.2).

On the other hand, observe that the separability is invariant under different choices of metrics on E generating the same topology.

Proof of Proposition 4.2. If E is totally bounded, then the unit ball is separable for the supremum norm, due to Proposition 4.1. Moreover, $\mathfrak{BL}_1(E)$ is a uniformly bounded family of Borel functions on E. Thus, since $\mathfrak{BL}_1(E)$ forms a universal Glivenko-Cantelli class, Theorem 1.3 combined with Corollary 1.4 in [169] applies, so that in particular we have

$$\sup\{|(m_n - \upsilon)f| : f \in \mathfrak{BL}_1(E)\} \to 0, \qquad \text{a.s. as } n \to +\infty. \tag{4.4}$$

The proof now concludes since (4.4) implies that $\gamma(m_n, \upsilon) \to 0$ a.s., as $n \to +\infty$, so that a.s. $m_n \to \upsilon$ in the weak topology, as $n \to +\infty$. $\square$

Topologies induced by families of functions. Suppose $\mathfrak{F}$ to be a class of $\mathscr{E}$-measurable functions on E and set $\mathfrak{M}_1^{\mathfrak{F}}(E) \triangleq \{\mu \in \mathfrak{M}_1(E) : \mu|f| < \infty\}$. It turns out to be convenient to understand the space $\mathfrak{M}_1^{\mathfrak{F}}(E)$ as a convex subspace of the product space $\mathbb{R}^{\mathfrak{F}}$. For this purpose, assume that $\mathfrak{F}$ separates the points of $\mathfrak{M}_1^{\mathfrak{F}}(E)$, i.e. given any couple $\mu, \nu \in \mathfrak{M}_1^{\mathfrak{F}}(E)$, one has $\mu = \nu$ if and only if $\mu f = \nu f$ for any $f \in \mathfrak{F}$. Thus, we can associate any $\mu \in \mathfrak{M}_1^{\mathfrak{F}}(E)$ to the linear map $f \in \mathfrak{F} \mapsto \mu f$. As a result, the space $\mathbb{R}^{\mathfrak{F}}$ induces a topological structure on $\mathfrak{M}_1^{\mathfrak{F}}(E)$ in a natural way.

Definition 4.1 ($\mathfrak{F}$-weak topology). We call $\mathfrak{F}$-*weak topology* the topology $\sigma(\mathfrak{M}_1^{\mathfrak{F}}(E), \mathfrak{F})$ on $\mathfrak{M}_1^{\mathfrak{F}}(E)$ inherited from the product topology defined on $\mathbb{R}^{\mathfrak{F}}$.

According to Definition 4.1, the $\mathfrak{F}$-weak topology is thus the projective topology on $\mathfrak{M}_1^{\mathfrak{F}}(E)$ defined by the linear forms on $\mathfrak{M}_1^{\mathfrak{F}}(E)$ belonging to $\mathfrak{F}$, i.e. the coarsest topology on $\mathfrak{M}_1^{\mathfrak{F}}(E)$ that renders continuous the maps $\mu \in \mathfrak{F} \mapsto \mu f$. It is worth noting that, if $\langle \mathfrak{ca}(E), \mathfrak{F} \rangle$ provides a dual pairing in the duality

$$\langle \mu, f \rangle = \mu f, \qquad \text{for any } \mu \in \mathfrak{ca}(E) \text{ and } f \in \mathfrak{F}, \tag{4.5}$$

where $\mathfrak{ca}(E)$ denotes the family of all signed measures of bounded variation on $(E, \mathscr{E})$, and $\mathfrak{M}_1^{\mathfrak{F}}(E)$ is a $\sigma(\mathfrak{ca}(E), \mathfrak{F})$-closed and convex subspace of $\mathfrak{ca}(E)$, the $\mathfrak{F}$-weak topology on $\mathfrak{M}_1^{\mathfrak{F}}(E)$ is nothing but the relativization of $\sigma(\mathfrak{ca}(E), \mathfrak{F})$ to $\mathfrak{M}_1^{\mathfrak{F}}(E)$. Hence, according to the previous arguments, we also say that $\sigma(\mathfrak{M}_1^{\mathfrak{F}}(E), \mathfrak{F})$ is *weakly generated* by $\mathfrak{F}$, or *generated by the duality* with $\mathfrak{F}$. Moreover, a sequence in $\mathfrak{M}_1^{\mathfrak{F}}(E)$ is $\mathfrak{F}$-*weak convergent*, if it converges with respect to the $\mathfrak{F}$-weak topology on $\mathfrak{M}_1^{\mathfrak{F}}(E)$.

Notice that the previous arguments do not depend of the actual structure of $\mathfrak{F}$, for which we just required to be some set of $\mathscr{E}$-measurable functions on E. Thus, let us further assume that $\mathfrak{F}$ is a linear space and $\|\cdot\|_{\mathfrak{F}}$ a norm properly defined on it. We denote by $\mathfrak{F}_1$ the unit ball in $\mathfrak{F}$, i.e. we set $\mathfrak{F}_1 \triangleq \{f \in \mathfrak{F}, \|f\|_{\mathfrak{F}} \leq 1\}$.

Given a sequence $\mu_0, \mu_1, \ldots$ in $\mathfrak{M}_1^{\psi}(E)$ such that $\sup\{|(\mu_n - \mu_0)f| : f \in \mathfrak{F}_1\} \to 0$, as $n \to +\infty$, then $\mu_n f \to \mu_0 f$, for any $f \in \mathfrak{F}$, and hence $\mu_n \to \mu_0$ in the $\mathfrak{F}$-weak topology, as $n \to +\infty$. Then, assuming that $\mathfrak{F}_1$ is a μ-Glivenko Cantelli class, one has that $m_n \to \mu$ a.s. in the $\mathfrak{F}$-weak topology, as $n \to +\infty$, whenever the variables $\xi_1, \xi_2, \ldots$ directing the empirical process $(m_n)_n$ are assumed to be independent with common law $\mu \in \mathfrak{M}_1^{\mathfrak{F}}(E)$. This fact naturally extends to the case when $\xi_1, \xi_2, \ldots$ are assumed to be stationary.

Proposition 4.3. *Assume that $\mathfrak{F}_1$ is uniformly bounded and equicontinuous, and $(\xi_n)_n$ is stationary, then $m_n \to \upsilon$ in the $\mathfrak{F}$-weak topology.*

Proof. Since $(\xi_n)_n$ is assumed to be stationary, Proposition 4.2 guarantees that $m_n \to \upsilon$ a.s. in the weak topology as $n \to +\infty$. Hence, since $\mathfrak{F}_1$ is assumed to be equicontinuous and uniformly bounded, the result follows from Corollary 11.3.4 in [63]. $\square$

Weak Topology. Notice that when setting $\mathfrak{F}$ to be the space $\mathfrak{BL}(E)$ of bounded and Lipschitz continuous functions on E, Proposition 4.3 boils down to Proposition 4.2. We now show that the same result is obtained by considering any space of bounded and uniformly continuous functions characterized by a determined modulus of continuity.

We shall write $\mathfrak{A}(E)$ to denote the family of bounded and uniformly continuous functions on E. Given a modulus of continuity κ, for any real valued function f, the *Lipschitz seminorm* related to κ is defined as $\|f\|_{\mathfrak{L},\kappa} \triangleq |f(x) - f(y)|/\kappa(d(x, y))$. Hence, let $\|\cdot\|_\kappa \triangleq \|\cdot\|_\infty + \|\cdot\|_{\mathfrak{L},\kappa}$ and $\mathfrak{A}_{\kappa,b}(E) \triangleq \{f \in \mathfrak{A}(E) : \|f\|_{\kappa,b} < \infty\}$ be the family of bounded and uniformly continuous functions that displays κ as modulus of continuity.

Notice that the family of laws $\mu \in \mathfrak{M}_1(E)$ for which $\mu|f| < \infty$ for any $f \in \mathfrak{A}_{\kappa,b}(E)$ indeed coincides with $\mathfrak{M}_1(E)$. We define the κ-weak topology to be the weak(-star) topology generated on $\mathfrak{M}_1(E)$ by the duality of $\mathfrak{A}_{\kappa,b}(E)$.

Lemma 4.1. *The space $\mathfrak{A}_{\kappa,b}(E)$ is an algebra of functions and $\|\cdot\|_\kappa$ is a norm on it.*

Proof. Clearly, $\mathfrak{A}_{\kappa,b}(E)$ is a linear space. Let $f, g \in \mathfrak{A}_{b,\kappa}(E)$. To prove that $\mathfrak{A}_{\kappa,b}(E)$ is indeed an algebra we have to show that $fg \in \mathfrak{A}_{\kappa,b}(E)$. Thus, notice that

$$\begin{aligned} |f(x)g(x) - f(y)g(y)| &\leq |f(x)(g(x) - g(y))| - |g(y)(f(x) - f(y))| \\ &\leq \{\|f\|_\infty \|g\|_{\mathfrak{L},\kappa} + \|g\|_\infty \|f\|_{\mathfrak{L},\kappa}\}\kappa(d(x,y)), \end{aligned} \tag{4.6}$$

hence $fg \in \mathfrak{A}_{\kappa,b}(E)$. To conclude, observe that $\|\cdot\|_\kappa$ is a norm since $\|\cdot\|_{\mathfrak{L},\kappa}$ is a pseudonorm and $\|\cdot\|_\infty$ is a norm. □

Notice that when the space E is assumed to be compact, the family $\mathfrak{A}(E)$ indeed coincides with $\mathfrak{C}(E)$. In particular, since $\mathfrak{A}_{\kappa,b}(E)$ is an algebra of functions according to Lemma 4.1, it contains the constant and separates points, i.e. given any distinct points $x, y \in E$ one has $g(x) \neq g(y)$ for some $\mathfrak{A}(E)$, we have that Stone-Weiertsrass theorem ([63], Theorem 2.4.11) applies, and the result in Proposition 4.4 boils down to the following.

Lemma 4.2. *If E is compact, then the space $\mathfrak{A}_{\kappa,b}(E)$ is uniformly dense in $\mathfrak{A}(E)$.*

Notice that the space $\mathfrak{A}(E)$, as well as $\mathfrak{A}_{\kappa,b}(E)$, actually depends on the specific metric d on E consistent with its topological structure. In particular, the specific metric d consistent with the topology defined on E may fail to be complete, since completeness is not a topological invariant. On the other hand, there always exists a continuous embedding $\imath : E \hookrightarrow \bar{E}$, where the completion $\bar{E}$ of E is metrized by the distance $\bar{d}$ satisfying $\bar{d}(\imath x, \imath y) = d(x,y)$, for any $x, y \in E$, (see e.g. [63], § 2.5).

Lemma 4.3. *Any $g \in \mathfrak{A}(E)$ admits a unique extension $\bar{g} \in \mathfrak{A}(\bar{E})$. Moreover, one has $\|g\|_\infty = \|\bar{g}\|_\infty$.*

Proof. For any $x \in \bar{E}$, define $\bar{g}(x) \triangleq \lim_n g(x_n)$, where $x_1, x_2, \ldots$ is some sequence in E satisfying $x_n \to x$ as $n \to +\infty$. Easily, the limit is well defined since $g(x_1), g(x_2), \ldots$ is a Cauchy sequence in $\bar{E}$ thanks to the uniformly continuity of g, and hence it converges since $\bar{E}$ is complete. Moreover, if $x_1^{(i)}, x_2^{(i)}, \ldots$ is a sequence in E such that $x_n^{(i)} \to +\infty$, for $i = 1, 2$, then one has that $\lim_n g(x_n^{(1)}) = \lim_n g(x_n^{(2)}) = \bar{g}(x)$ and hence the previous definition does not depend on the choice of the sequence $x_1, x_2, ...$ in E.

The function $\bar{g}$ turns out to be uniformly continuous on $\bar{E}$. Indeed, for any $x, y \in \bar{E}$ let $x_1, x_2, ...$ and $y_1, y_2, ...$ be two sequences in E such that $x_n \to x$ and $y_n \to y$, for $n \to +\infty$. Then, one has that

$$\bar{d}(\bar{g}(x), \bar{g}(y)) \leq \bar{d}(\bar{g}(x), g(x_n)) + d(g(x_n), g(y_n)) + \bar{d}(g(y_n), \bar{g}(x)), \tag{4.7}$$

and hence the uniformly continuity of $\bar{g}$ follows form the uniformly continuity of g, provided that n is large enough.

The function $\bar{g}$ is clearly an extension of g and it is unique. Indeed, if $\bar{h}$ is another continuous extension of g in $\bar{E}$, one should have $\bar{h}(x_n) \to \bar{h}(x)$, for any sequence $x_1, x_2, ...$ in E such that $x_n \to x \in \bar{E}$, for $n \to +\infty$. On the other hand, one has $\bar{h}(x_n) = g(x_n) \to \bar{g}(x)$, for $n \to +\infty$, and hence $\bar{h}(x) = \bar{g}(x)$ must hold.

Finally, the identity $\|g\|_\infty = \|\bar{g}\|_\infty$ easily follows from standard continuity arguments. □

The following proposition extends the result in Lemma 4.2.

Proposition 4.4. *The space $\mathfrak{A}_{\kappa,b}(E)$ is uniformly dense in $\mathfrak{A}(E)$.*

Recall that it is not restrictive to assume that E is totally bounded. Indeed, among all the metrics on E that are consist with its topological structure, there exists one that is totally bounded, (cf. [63], Theorem 2.8.2).

Proof of Proposition 4.4. Notice that since E is totally bounded, its completion $\bar{E}$ appears to be compact. Thus, let $g \in \mathfrak{A}(E)$ and define $g \in \mathfrak{A}(\bar{E})$ to be its unique extension provided by Lemma 4.3. According to Lemma 4.2, there exists a sequence $(\bar{g}_n)_n$ in $\mathfrak{A}_{\kappa,b}(\bar{E})$ such that $\|\bar{g} - \bar{g}_n\|_\infty \to 0$, as $n \to +\infty$. According to the definition of $\bar{d}$ we have that for any $n \geq 1$ the restriction $g_n \triangleq \bar{g}_{n|_E}$ of $\bar{g}_n$ to E belongs to $\mathfrak{A}_{\kappa,b}(E)$ and $\|g - g_n\|_\infty = \|\bar{g} - \bar{g}_n\|_\infty \to 0$, as $n \to +\infty$. □

Since any uniformly dense subset of $\mathfrak{A}(E)$ generates the weak topology on $\mathfrak{M}_1(E)$ by duality (cf. [6], Theorem 15.2), Proposition 4.4 leads to the following result.

Corollary 4.1. *The κ-weak topology coincides with the weak topology on $\mathfrak{M}_1(E)$.*

4.2 ψ-WEAK TOPOLOGY OF MEASURES

In many circumstances, the weak topology of measures turns out to be too coarse. One of the typical problems is that some common functionals on $\mathfrak{M}_1(E)$ endowed with the weak topology of measures, like for instance the mean, fail to be continuous. On the other hand, since the weak topology looks at the body of the distributions, the tail part is completely neglected. This means that two laws may be close with respect to some metric for the weak convergence and at the same time they may considerably deviate on their tails. Indeed, recall that the family of Borel measure with bounded support is weakly dense in $\mathfrak{M}_1(E)$.

Throughout this section we generalize the weak topology of measures by introducing a stronger notion of topology.

The ψ-weak topology. Let ψ be a continuous function defined on E, satisfying $\psi \geq 1$. Any function of this kind is termed a *gauge function* on E. Thus, consider the space $\mathfrak{C}_\psi(E)$ of continuous functions given by

$$\mathfrak{C}_\psi(E) \triangleq \{f \in \mathfrak{C}(E) : \|f/\psi\|_\infty < \infty\}.$$

The function ψ plays the role of penalizing the displacement of the distributions at their tails. On the other hand, the space $\mathfrak{C}_\psi(E)$ consists of those continuous functions whose growth is controlled by ψ. Clearly, such a space boils down to $\mathfrak{C}_b(E)$ in the special case when $\psi \equiv 1$ or simply when ψ is bounded above. Thus, consider the subspace $\mathfrak{M}_1^\psi(E)$ of $\mathfrak{M}_1(E)$ given by

$$\mathfrak{M}_1^\psi(E) \triangleq \{\mu \in \mathfrak{M}_1(E) : \mu\psi < +\infty\}.$$

Since $\mathfrak{M}_1^\psi(E)$ is contained in $\mathfrak{M}_1(E)$, it inherits the topological structure provided by the relativization of the weak topology on $\mathfrak{M}_1(E)$. This relative weak topology is actually $\sigma(\mathfrak{M}_1^\psi(E), \mathfrak{C}_b(E))$, the coarsest topology so that for each $f \in \mathfrak{C}_b(E)$, the mapping $\mu \in \mathfrak{M}_1^\psi(E) \mapsto \mu f$ is continuous,(cf. [6], Lemma 2.53). Moreover, the space $\mathfrak{M}_1^\psi(E)$ turns out to be separable when endowed with the topology $\sigma(\mathfrak{M}_1^\psi(E), \mathfrak{C}_b(E))$. Indeed, notice that if $e_1, e_2, \ldots$ is a sequence dense in E, then the family of convex combinations (with rational weights) of δ_{e_j} is contained in $\mathfrak{M}_1^\psi(E)$ and it is dense in $\mathfrak{M}_1(E)$, with respect to Prohorov distance.

On the other hand, we consider on $\mathfrak{M}_1^\psi(E)$ the topology defined by the duality with $\mathfrak{C}_\psi(E)$ as follows.

Definition 4.2 (ψ-weak topology). The *ψ-weak topology* $\sigma(\mathfrak{M}_1^\psi(E), \mathfrak{C}_\psi(E))$ is the coarsest topology on $\mathfrak{M}_1^\psi(E)$ that renders continuous the maps $\mu \in \mathfrak{M}_1^\psi(E) \mapsto \mu f$, varying $f \in \mathfrak{C}_\psi(E)$.

Generally, the ψ-weak topology appears to be finer than the relative weak topology $\sigma(\mathfrak{M}_1^\psi(E), \mathfrak{C}_b(E))$. These two topological structures indeed coincide only when $\psi \equiv 1$ or simply ψ is bounded above. Moreover, when endowed with $\sigma(\mathfrak{M}_1^\psi(E), \mathfrak{C}_\psi(E))$, one has that $\mathfrak{M}_1^\psi(E)$ acquires the structure of Polish space (cf. [83], Corollary A.45). More precisely, the ψ-weak topology may be metrized by the distance d_ψ defined as follow, (cf. [114], Lemma 3.4),

$$d_\psi(\mu,\nu) \triangleq \pi(\mu,\nu) + |(\mu-\nu)\psi|, \qquad \text{for any } \mu,\nu \in \mathfrak{M}_1^\psi(E). \tag{4.8}$$

When looking at the metric defined in (4.8), it is easily to be realized that given a sequence $\mu_0, \mu_1, \ldots$ in $\mathfrak{M}_1^\psi(E)$, one has that $\mu_n \to \mu_0$ in the ψ-weak topology as $n \to +\infty$ if and only if $\mu_n \to \mu_0$ in the relative weak topology $\sigma(\mathfrak{M}_1^\psi(E), \mathfrak{C}_b(E))$ and $\mu_n\psi \to \mu_0\psi$ as $n \to +\infty$.

Memorability Issues. Throughout, we discuss the measurable structure on $\mathfrak{M}_1^\psi(E)$ induced by the ψ-weak topology. Thus, define $\mathscr{M}^\psi$ to be the Borel σ-algebra generated by the ψ-weak topology $\sigma(\mathfrak{M}_1^\psi(E), \mathfrak{C}_\psi(E))$ on $\mathfrak{M}_1^\psi(E)$. The following result characterizes $\mathscr{M}^\psi$ in terms of the relative weak topology.

Lemma 4.4. *The σ-algebra $\mathscr{M}^\psi$ is generated by the relative weak topology on $\mathfrak{M}_1^\psi(E)$, i.e. $\sigma(\mathfrak{M}_1^\psi(E), \mathfrak{C}_b(E))$.*

Recall that a σ-algebra is said to be (i) *countably generated* if it is generated by a countable family of sets and (ii) *countably separated* if it admits an a countable family of sets separating points. Moreover, a measurable space is said to be *standard* if it is Borel-isomorphic to a Polish space.

Proof of Lemma 4.4. Let us define $\mathscr{B}^\psi$ to be the Borel σ-algebra associated to the relative weak topology $\sigma(\mathfrak{M}_1^\psi(E), \mathfrak{C}_b(E))$. Since the ψ-weak topology is finer than the weak topology on $\mathfrak{M}_1^\psi(E)$, one has $\mathscr{B}^\psi \subseteq \mathscr{M}^\psi$. Since $\mathfrak{M}_1^\psi(E)$ is separable when endowed with the relative weak topology, the σ-algebra $\mathscr{B}^\psi$ is countably generated and countably separated. Indeed, any countable base $\mathfrak{B}^\psi$ of open sets in $\sigma(\mathfrak{M}_1^\psi(E), \mathfrak{C}_b(E))$ generates the σ-algebra $\mathscr{B}^\psi$ and separates points, (cf. [21], §6.5).

On the other hand, the space $(\mathfrak{M}_1^\psi(E), \mathscr{M}^\psi)$ is standard, and thus, the σ-algebra $\mathscr{M}^\psi$ coincides with $\sigma(\mathfrak{B}^\psi)$, thanks to Theorem 3.3 in [129]. □

Let us now define $\mathscr{M}$ to be the Borel σ-algebra generated by the weak topology $\sigma(\mathfrak{M}_1(E), \mathfrak{C}_b(E))$ on $\mathfrak{M}_1(E)$. It is worth to recall that $\mathscr{M}$ admits the useful characterization as the σ-algebra on $\mathfrak{M}_1(E)$ generated by the projection maps $\pi_B : \mu \in \mathfrak{M}_1(E) \mapsto \mu(B)$, letting B vary in $\mathscr{E}$, (cf. [90], Proposition 2.2.2). When looking at Lemma 4.4, one might be tempted to attend that a similar characterization indeed applies also when considering the σ-algebra $\mathscr{M}^\psi$.

Proposition 4.5. *The Borel σ-algebra $\mathscr{M}^\psi$ is generated by the projections $\pi_B : \mu \mapsto \mu(B)$, defined for $\mu \in \mathfrak{M}_1^\psi(E)$, letting B vary in $\mathscr{E}$.*

In order to prove the statement in Proposition 4.5, the following result appears to be useful.

Lemma 4.5. *Let $\mathfrak{H}$ be a family of functions defined on a set H and taking values in a measurable space $(G, \mathscr{G})$. Let ϕ be a H-valued map defined on some set H_0, then $\phi^{-1}(\sigma(\mathfrak{H})) = \sigma(\mathfrak{H} \circ \phi)$ on H_0, where $\mathfrak{H} \circ \phi \triangleq \{h \circ \phi : h \in \mathfrak{H}\}$.*

Proof of Lemma 4.5. First of all, note that $\phi^{-1}(\sigma(\mathfrak{H}))$ is a σ-algebra on H_0, since the map ϕ^{-1} preserves all the set operations. Thus, the inclusion $\sigma(\mathfrak{H} \circ \phi) \subseteq \phi^{-1}(\sigma(\mathfrak{H}))$ is immediate, since $h \circ \phi$ is $\phi^{-1}(\sigma(\mathfrak{H}))$-measurable for any $h \in \mathfrak{H}$.

Let now $\mathscr{H}_0$ be a σ-algebra on H_0 with respect to which $h \circ \phi$ is $(\mathscr{H}_0, \mathscr{G})$-measurable, for any $h \in \mathfrak{H}$. Clearly $\phi^{-1}(\sigma(\mathfrak{H})) \subseteq \mathscr{H}_0$. Thus, the proof concludes by considering $\mathscr{H}_0 = \sigma(\mathfrak{H} \circ \phi)$. □

Proof of Proposition 4.5. Let $\phi : \mathfrak{M}_1^{\psi}(E) \hookrightarrow \mathfrak{M}_1(E)$ be the inclusion of $\mathfrak{M}_1^{\psi}(E)$ into $\mathfrak{M}_1(E)$ and define $\mathfrak{H}$ to be the family consisting of the projection maps $\pi_B : \mu \in \mathfrak{M}_1(E) \to \mu(B)$, letting B vary in $\mathscr{E}$. The family $\mathfrak{H} \circ \phi \triangleq \{\pi_B \circ \phi : B \in \mathscr{E}\}$ consists of the projections defined on $\mathfrak{M}_1^{\psi}(E)$.

Let us define $\mathscr{B}^{\psi}$ to be the Borel σ-algebra generated by the relative weak topology $\sigma(\mathfrak{M}_1^{\psi}(E), \mathfrak{C}_b(E))$. The equality $\mathscr{B}^{\psi} = \phi^{-1}(\sigma(\mathfrak{H}))$ holds true, since $\sigma(\mathfrak{H}) = \mathscr{M}$ and $\mathscr{B}^{\psi} = \phi^{-1}(\mathscr{M})$. Hence, applying Lemma 4.5, one deduces that $\mathscr{B}^{\psi} = \sigma(\mathfrak{H} \circ \phi)$. The stated result now follows from Lemma 4.4. □

4.3 STRONG CONSISTENCY OF ESTIMATORS

Later on, we fix a gauge function ψ on E. Notice that $m_n \in \mathfrak{M}_1^{\psi}(E)$ a.s. for any $n \geq 1$. From Proposition 4.5 the σ-algebra $\mathscr{M}^{\psi}$ is generated by the projection mappings $\pi_B : \mu \in \mathfrak{M}_1^{\psi}(E) \mapsto \mu(B)$, varying $B \in \mathscr{E}$, then it is possible to look at the process $(m_n)_n$ as a random sequence in $(\mathfrak{M}_1^{\psi}(E), \mathscr{M}^{\psi})$.

Let T be a Polish space endowed with the related Borel σ-algebra $\mathscr{T}$. We shall denote by d_T the generic metric consistent with its topological structure. Throughout, we consider a generic functional τ defined on $\mathfrak{M}_1^{\psi}(E)$ and taking values in T. More precisely, any $(\mathscr{M}^{\psi}, \mathscr{T})$-measurable functional is termed a *statistic* over $\mathfrak{M}_1^{\psi}(E)$, and T its *action domain*. For the sake of simplicity, we say that τ is ψ-continuous when it is continuous with respect to the ψ-weak topology on $\mathfrak{M}_1^{\psi}(E)$ and the topology defined on T respectively. Besides, the statistic τ is said to be ψ-uniformly continuous if for any $\varepsilon > 0$ there exists $\delta(\varepsilon) > 0$ such that $d_t(\tau(\mu_1), \tau(\mu_2)) < \varepsilon$ if $d_{\psi}(\mu_1, \mu_2) < \delta(\varepsilon)$.

Notice that, if τ is ψ-continuous, then any element of the random sequence $\tau_1, \tau_2, \ldots$ in T obtained by setting $\tau_n \triangleq \tau(m_n)$, for any $n \geq 1$, turns out to be $(\mathscr{F}, \mathscr{T})$-measurable. In particular, we refer to $\tau_1, \tau_2, \ldots$ as the sequence of *estimators* induced by τ.

DEFINITION 4.3 (Strong Consistency). If $\tau_n \to \tau(\upsilon)$ a.s as $n \to +\infty$, we say that the random sequence $(\tau_n)_n$ provides a family of *strongly consistent* estimators for $\tau(\upsilon)$, or simply that the statistic τ is *strongly consistent* with respect to the random measure υ.

Assuming that ξ is stationary, Proposition 4.2 states that $m_n \to \upsilon$ in the weak topology as $n \to +\infty$. Hence, the strong consistency of the estimators τ_n may be properly achieved by natural arguments involving the continuity of the statistic τ. Proposition 4.2 may be generalized when considering the ψ-weak topology in a natural way.

Theorem 4.1. *If ξ is stationary and such that $\mathscr{L}(\xi_1) \in \mathfrak{M}_1^{\psi}(E)$, then $m_n \to \upsilon$ a.s. in the ψ-weak topology, as $n \to +\infty$.*

Proof. According to Proposition 4.2, we have that $\mathbb{P}$-almost surely $m_n \to \upsilon$ in the relative weak topology, as $n \to +\infty$. Thus, it remains to show that

$$|(m_n - \upsilon)\psi| \to 0, \qquad \text{a.s. as } n \to +\infty.$$

We apply now Theorem 2.8. Using the notation therein, we consider as transformation $T \triangleq \Sigma$, the shift operator on $E^{\mathbb{N}}$, and as measurable function $f \triangleq \psi \circ \xi_1$. Notice that the shift operator T preserves the measure $\mathbb{P}$ since ξ is assumed to be stationary. If we write μ for the common law $\mu = \mathscr{L}(\xi_1)$, then a change of variables gives that

$$\int_\Omega f d\mathbb{P} = \int_\Omega \psi(\xi_1) d\mathbb{P} = \int_E \psi(x) \mathbb{P} \circ \xi_1^{-1}(dx) = \int_E \psi d\mu,$$

since by hypothesis $\mathscr{L}(\xi_1) = \mu \in \mathfrak{M}_1^{\psi}(E)$, we have $\int_E \psi \, d\mu = \mu\psi < +\infty$ and therefore $f \in L^1(\Omega, \mathscr{F}, \mathbb{P})$. Hence, we conclude that

$$m_n\psi \to \mathbb{E}[\psi(\xi_1)|\xi^{-1}\mathscr{I}], \qquad \text{a.s. as } n \to +\infty.$$

Finally, according to the *disintegration theorem* (cf. [107], Theorem 5.4), we may recast the limit variable and write

$$\mathbb{E}[\psi(\xi_1)|\xi^{-1}\mathscr{I}] = \upsilon\psi, \qquad \text{a.s.}$$

$\square$

Notice that since the stationary structure of the sequence ξ implies that $\mathscr{L}(\xi_n) = \mathscr{L}(\xi_1)$ for any $n \geq 2$, if in addition ξ is ergodic, i.e. its distribution on $(E^{\mathbb{N}}, \mathscr{E}^{\mathbb{N}})$ is ergodic with respect to the shift operator Σ on $E^{\mathbb{N}}$, the asymptotic result stated by Theorem 4.1 boils down to $\tau_n \to \tau(\mu)$, $\mathbb{P}$-almost certainty as $n \to +\infty$, where we set $\mu \triangleq \mathscr{L}(\xi_1)$. Indeed the probability measure $\mathbb{P}$ turns out to be trivial on $\mathscr{I}_\Sigma$ in such a case.

The next result is an immediate consequence of Theorem 4.1.

Corollary 4.2 (*Strong Consistency*)**.** *If ξ is stationary and $\tau : \mathfrak{M}_1^\psi(E) \to T$ is ψ-continuous, then the sequence of estimators $(\tau_n)_n$ is strongly consistent for $\tau(\upsilon)$.*

4.4 ROBUSTNESS IN A PERTURBATIVE SETTING

In this section, we introduce a refined notion of robustness in terms of the canonical random measure associated to a given stationary sequence in E. As a result, a similar characterization as the one proposed in Theorem 3.8 is recovered in terms of the modulus of continuity of the statistic τ.

Robustness. Throughout, given any fixed $\mathscr{F}$-measurable endomorphism θ properly defined over Ω, i.e. a $(\mathscr{F}, \mathscr{F})$-measurable map $\theta : \Omega \to \Omega$, we write $\mathbb{P}_\theta \triangleq \mathbb{P} \circ \theta^{-1}$ to denote the measure given by the image of $\mathbb{P}$ under the action of θ. Moreover, we shall assume that the measures $\mathbb{P}_\theta$ and $\mathbb{P}$ are equivalent. In this case, we say that the measure $\mathbb{P}$ is *quasi-invariant* under the action of θ, and we shortly write $\mathbb{P}_\theta \simeq \mathbb{P}$.

Recall that, in the case when $\mathscr{L}(\xi_1) \in \mathfrak{M}_1^\psi(E)$, the random measure υ may be understood as a random element of $(\mathfrak{M}_1^\psi(E), \mathscr{M}^\psi)$. Thus, denote by $\upsilon \circ \theta$ the probability kernel obtained by setting $(\upsilon \circ \theta)(\omega, B) \triangleq \upsilon(\theta(\omega), B)$, for any $\omega \in \Omega$ and $B \in \mathscr{E}$.

Since the metric d_ψ is clearly $\mathscr{M}^\psi \otimes \mathscr{M}^\psi$-measurable, we are allowed to define

$$\lambda_\theta(\alpha) \triangleq \mathbb{P}\{d_\psi(\upsilon, \upsilon \circ \theta) > \alpha\}, \qquad \text{for any } \alpha \geq 0. \tag{4.9}$$

It is worth to be noticed that the map λ_θ defined in (4.9) depends also on the probability measure $\mathbb{P}$. For simplicity, we do not explicit this dependence and we shall use the notation above.

Assume that the statistic τ is uniformly ψ-continuous and that κ is a modulus of continuity of τ, i.e. a continuous strictly increasing map $[0, +\infty] \to [0, +\infty]$ such that

$$d_\tau(\tau(\mu_1), \tau(\mu_2)) \leq \kappa(d_\psi(\mu_1, \mu_2)), \qquad \text{for any } \mu_1, \mu_2 \in \mathfrak{M}_1^\psi(E).$$

Note that, since $\lambda_\theta(\alpha) < \kappa(\alpha)$ for α large enough, we are allowed to set

$$\|\theta\|_{\upsilon,\kappa} \triangleq \inf\{\alpha > 0 : \lambda_\theta(\alpha) < \kappa(\alpha)\}. \tag{4.10}$$

Lemma 4.6. *If τ is uniformly ψ-continuous and it admits κ as modulus of continuity, then $\pi(\mathbb{P} \circ \tau(\upsilon)^{-1}, \mathbb{P}_\theta \circ \tau(\upsilon)^{-1}) \leq \kappa(\|\theta\|_{\upsilon,\kappa})$.*

Proof. Let $C \in \mathscr{T}$ and fix $\alpha > 0$ such that $\lambda_\theta(\alpha) < \kappa(\alpha)$. Since τ is ψ-continuous, then $\tau^{-1}(C) \in \mathscr{M}^\psi$. In particular, for any $A \in \mathscr{M}^\psi$, we denote by $A^\varepsilon \triangleq \{\mu \in \mathfrak{M}_1^\psi(E) : d_\psi(\mu, \nu) \leq \varepsilon, \text{ for some } \nu \in A\}$ the ε-hull of A defined in terms of the metric d_ψ.

Notice that $[\tau^{-1}(C)]^\alpha \subseteq \tau^{-1}(C^{\kappa(\alpha)})$ in $\mathfrak{M}_1^\psi(E)$, since τ is uniformly ψ-continuous and admits κ as modulus of continuity, where the $\kappa(\alpha)$-hull $C^{\kappa(\alpha)}$ of C is defined in terms of the metric d_ψ. Hence, $\upsilon \circ \theta \in [\tau^{-1}(C)]^\alpha$ implies $\upsilon \circ \theta \in \tau^{-1}(C^{\kappa(\alpha)})$, and in particular one has that $\mathbb{P}\{\upsilon \circ \theta \in [\tau^{-1}(C)]^\alpha\} \leq \mathbb{P} \circ \tau(\upsilon \circ \theta)^{-1}(C^{\kappa(\alpha)})$. Thus,

$$\begin{aligned} \mathbb{P} \circ \tau(\upsilon)^{-1}(C) &\leq \mathbb{P}\{d_\psi(\upsilon, \upsilon \circ \theta) > \alpha\} + \mathbb{P}\{\upsilon \circ \theta \in [\tau^{-1}(C)]^\alpha\} \\ &\leq \kappa(\alpha) + \mathbb{P} \circ \tau(\upsilon \circ \theta)^{-1}\Big(C^{\kappa(\alpha)}\Big). \end{aligned}$$

Then, since the choice of $C \in \mathscr{T}$ is arbitrary, one has that

$$\pi\Big(\mathbb{P} \circ \tau(\upsilon)^{-1}, \mathbb{P} \circ \tau(\upsilon \circ \theta)^{-1}\Big) \leq \kappa(\alpha).$$

Hence, since $\mathbb{P}_\theta \circ \tau(\upsilon)^{-1} = \mathbb{P} \circ \tau(\upsilon \circ \theta)^{-1}$, the proof is concluded by letting α tend to $\|\theta\|_{\upsilon,\kappa}$, while invoking the continuity of κ. $\square$

Notice that, in the special case when κ is defined as the identity over $(0, +\infty)$, one has that (4.10) boils down to the Ky Fan type distance $\chi_\psi(\upsilon, \upsilon \circ \theta)$ relative to d_ψ between υ and $\upsilon \circ \theta$ given by

$$\chi_\psi(\upsilon, \upsilon \circ \theta) \triangleq \inf\{\varepsilon > 0 : \mathbb{P}(d_\psi(\upsilon, \upsilon \circ \theta) > \varepsilon) \leq \varepsilon\}. \tag{4.11}$$

Here, the random measures υ and $\upsilon \circ \theta$ are to be understood as random elements in $(\mathfrak{M}_1^\psi(E), \mathscr{M}^\psi)$. In particular, when looking at Lemma 4.6, this is the case when τ is indeed assumed to be a contraction.

Theorem 4.2 (Robustness). *Let ξ be stationary and let $\mathbb{P}$ be quasi-invariant under θ. If τ is uniformly ψ-continuous and it admits κ as modulus of continuity, then*

$$\limsup_{n \geq 1} \ \pi(\mathbb{P} \circ \tau_n^{-1}, \mathbb{P}_\theta \circ \tau_n^{-1}) \leq \kappa(\|\theta\|_{\upsilon,\kappa}). \tag{4.12}$$

Proof. By the triangle inequality we have,

$$\begin{aligned} \pi(\mathbb{P} \circ \tau_n^{-1}, \mathbb{P}_\theta \circ \tau_n^{-1}) \leq\ & \pi(\mathbb{P} \circ \tau_n^{-1}, \mathbb{P} \circ \tau(\upsilon)^{-1}) \\ & + \pi(\mathbb{P} \circ \tau(\upsilon)^{-1}, \mathbb{P}_\theta \circ \tau(\upsilon)^{-1}) \\ & + \pi(\mathbb{P}_\theta \circ \tau_n^{-1}, \mathbb{P}_\theta \circ \tau(\upsilon)^{-1}). \end{aligned} \tag{4.13}$$

Since τ is uniformly ψ-continuous and admits κ as modulus of continuity, Lemma 4.6 applies. Thus, from inequality (4.13) we deduce that

$$\begin{aligned} \limsup_{n \geq 1} \ \pi(\mathbb{P} \circ \tau_n^{-1}, \mathbb{P}_\theta \circ \tau_n^{-1}) \leq\ & \kappa(\|\theta\|_{\upsilon,\kappa}) \\ & + \limsup_{n \geq 1} \{\pi(\mathbb{P} \circ \tau_n^{-1}, \mathbb{P} \circ \tau(\upsilon)^{-1}) + \pi(\mathbb{P}_\theta \circ \tau_n^{-1}, \mathbb{P}_\theta \circ \tau(\upsilon)^{-1})\}. \end{aligned} \tag{4.14}$$

On the other hand, ξ is assumed to be stationary and τ is ψ-continuous. Then, the result described in Corollary 4.2 guarantees that $\mathbb{P}$-almost surely $\tau_n \to \tau(\upsilon)$, as $n \to +\infty$, and hence also $\mathbb{P}_\theta$-almost surely, as $\mathbb{P}$ is assumed to be quasi-invariant under θ. Thus,

$$\limsup_{n \geq 1} \{\pi(\mathbb{P} \circ \tau_n^{-1}, \mathbb{P} \circ \tau(\upsilon)^{-1}) + \pi(\mathbb{P}_\theta \circ \tau_n^{-1}, \mathbb{P}_\theta \circ \tau(\upsilon)^{-1})\} = 0.$$

$\square$

Since according to our setup, any path of the process ξ, and hence any element of the sample space Ω, models the collected values, we may understand the action of the endomorphism θ as a specific perturbation of the dataset. In this respect, the function λ_θ defined in (4.9) gauges the impact of such a perturbation in terms of the random measure υ. In particular, note that $\|\theta\|_{\upsilon,\kappa} = 0$, when θ is chosen to be the identity over Ω, or more generally when the action of θ does not affect the distribution of the random measure υ. It is easy to realize that, when the perturbation procedure encoded by the action of the map θ does not change significantly the random measure υ in the stochastic sense provided by (4.9), then one should expect $\|\theta\|_{\upsilon,\kappa}$ to be small. In particular, this form of continuity is properly assessed in terms of κ. Indeed, in the particular case when τ admits κ as modulus of continuity, Theorem 4.2 guarantees that small perturbations at the level of the dataset only result in small perturbations in terms of the asymptotic law of the family of estimators associate to the statistic τ. In particular, the impact of the generic perturbation is precisely gauged by the relation described in (4.12).

Bayesian nonparametrics. At the level of the statistical inference, Theorem 4.2 admits also a natural interpretation in terms of bayesian analysis. In this respect, Corollary 4.2, as well as Theorem 4.2, still remains valid when the sequence of projections $\xi_1, \xi_2, \ldots$ is assumed to be exchangeable, since any exchangeable and numerable random sequence is always stationary, (cf. [106], Proposition 2.2).

On the other hand, when dealing with numerable exchangeable random sequences, we get that $\xi^{-1}\mathscr{I} = \sigma(\upsilon)$ a.s. thanks to Theorem 2.12. In addition, the σ-algebra $\sigma(\upsilon)$ turns out to be $\mathbb{P}$-trivial in the independence setup. In this respect, we are allowed to recast the limit random variable in Theorem 4.1 by writing $\upsilon = \mathbb{P}[\xi_1 \in \cdot\,|\upsilon]$, where the equality is intended in the $\mathbb{P}$-almost surely sense. On the other hand, according to Theorem 2.11, when dealing with a numerable random sequence $\xi = (\xi_1, \xi_2, \ldots)$ in E, the notion of exchangeability coincides with a conditional form of independence, i.e. one has that $\mathbb{P}[\xi \in \cdot\,|\upsilon] = \upsilon^{\mathbb{N}}$ a.s.

Exchangeability provides the main pillar of the Bayesian approach to the inferential analysis. More precisely, when dealing with the non parametric setup, the law induced by the random measure υ over the space $(\mathfrak{M}_1^\psi(E), \mathscr{M}^\psi)$ may be regarded as the *prior distribution* of the statistical model

$$\xi_1, \xi_2, ...|\upsilon \sim_{iid} \upsilon,$$

where the latter form of independence is to be understood in terms of de Finetti's theorem. The prior distribution is the main pillar of Bayesian statistics, and it represents the subjective knowledge before the experiment is conducted.
According to such a formulation, Theorem 4.2 may be regarded as a form of stability obtained when the prior distribution of the model is forced to change in such a way that the quasi-invariance of the measure $\mathbb{P}$ is guaranteed.

Robustness and back-testing. Here and in the sequel, a scoring function is any map $S : T \times E \to \mathbb{R}$ such that the function $x \in E \mapsto S(t, x)$ is integrable with respect to any distribution $\mu \in \mathfrak{M}_1^\psi(E)$, for any $t \in T$. The statistic τ is termed *elicitable* if there exists a scoring function $S : T \times E \to [0, +\infty)$ such that

$$\int_E S(\tau(\mu), x)\mu(dx) \leq \int_E S(t, x)\mu(dx), \qquad \text{for any } \mu \in \mathfrak{M}_1^\psi(E) \text{ and } t \in T. \tag{4.15}$$

and the identity in (4.15) holds if and only if $\tau(\mu) = t$.

The notion of elicitability provides a widely discussed aspect in evaluating point forecasts, since it may be regarded as the dual form of optimality in the point forecasting framework. As a consequence, many authors believe that elicitability provides the natural tool in order to make back-testing. For background see for instance [81, 93, 142, 189].

Throughout, we shall assume that τ is uniformly continuous with respect to the functional $(\mu, \nu) \mapsto \tilde{S}(\mu, \nu) \triangleq \int_E S(\tau(\mu), x)\nu(dx)$ in the sense that

$$d_T(\tau(\mu), \tau(\nu)) \leq \kappa(\tilde{S}(\mu, \nu)), \qquad \text{for any } \mu, \nu \in \mathfrak{M}_1^\psi(E), \tag{4.16}$$

for some non-negative continuous and increasing function κ vanishing at zero.

Recall that $\|\theta\|_{\mathbb{P},\kappa}$ as defined in (4.10) implicitly depends on the metric d_ψ. In a similar way, if $\tilde{S}$ is $(\mathscr{M}^\psi \otimes \mathscr{M}^\psi)$-measurable, we may define

$$\|\theta\|^{(1)}_{\mathbb{P},\kappa} \triangleq \inf\{\alpha > 0 : \mathbb{P}\{\tilde{S}(\upsilon, \upsilon \circ \theta) > \alpha\} < \kappa(\alpha)\}. \tag{4.17}$$

Hence, under condition (4.16) a similar estimate as provided in Lemma 4.6 may be assessed in terms of (4.17). As a result, the arguments in the proof of Theorem 4.2 still remain valid and give the following result.

Corollary 4.3. *Assume that the condition (4.16) holds true. If τ is assumed to be ψ-continuous, ξ is stationary and $\mathbb{P}$ is quasi-invariant under θ, then*

$$\limsup_{n\geq 1} \pi(\mathbb{P}\circ\tau_n^{-1}, \mathbb{P}_\theta \circ \tau_n^{-1}) \leq \kappa(\|\theta\|^{(1)}_{\mathbb{P},\kappa}).$$

As an example, when considering E and T to be the real line endowed with the Euclidean metric and ψ the identity, if $\tau : \mu \in \mathfrak{M}_1^\psi(E) \mapsto \int_{\mathbb{R}} x\mu(dx)$ defines the mean and $S(x,y) \triangleq (x-y)^2$, for any $(x,y) \in \mathbb{R}^2$, then condition (4.16) is guaranteed when for instance $\kappa(z) \triangleq \sqrt{z}$, for any $z \geq 0$.

Also, observe also that when ψ is strictly increasing and $\tau(\mu)$ is defined as the α-quantile of the law $\mu \in \mathfrak{M}_1^\psi(E)$, for some fixed $\alpha \in (0,1)$, and the related scoring function is given by $S(x,y) \triangleq (\mathbb{1}_{\{x\geq y\}} - \alpha)(\psi(x) - \psi(y))$, (see, e.g., Theorem 3.3 in [93]), condition (4.16) fails for any κ.

Part II

Portfolio Representation in Life Insurance

CHAPTER 5

Stochastic Integration in Banach Spaces

In the last decades, many results of stochastic integration in Hilbert Spaces have been generalized to a Banach space framework. A theory of stochastic integration for processes taking values in Banach spaces has been initially proposed by considering spaces with martingale type 2. In this respect, we refer to the works developed by Dettweiler [60, 61], Neidhardt [139] and Ondreját [141]. Some of the cited authors developed the theory of stochastic integration in the so-called 2-uniformly smooth Banach spaces. Besides, Pisier [145] proved that the martingale type 2 property is indeed equivalent to the 2-uniform smoothness. Brzeźniak [28, 29, 30, 27] further carried on the stochastic integrals of Neidhardt and Dettweiler by providing some applications to the theory of stochastic partial differential equations.

On a different line, Garling [88] and McConnell [134] exploited the probabilistic definition of the so-called UMD Banach spaces introduced by Burkholder [32] and Bourgain [23] to propose a theory of stochastic calculus in this class of spaces. The ideas proposed by Garling and McConnell have been extensively broadened by Van Neerven, Veraar and Weis [173, 174, 177]. We refer to [170, 176, 175] and [31] for the application of the theory of stochastic integration in UMD Banach spaces to the study of stochastic partial differential equations.

Motivated by these developments, some authors extended many of the standard notions of Malliavin calculus to a Banach space framework. We mainly refer to the work by Maas and van Neerven [128], in which a Clark-Ocone representation formula is presented when dealing with random variables taking values in a UMD Banach space. Other extensions of the Clark-Ocone formula may be found within the works of Mayer-Wolf and Zakai [131, 132], Osswald [143] and Faria, Oliveira and Streit [74].

Thorough, we shall fix a Banach space E and a separable Hilbert space H. The topological dual of E is written as E^*, while we shall always identify H with its topological dual H^* via the Riesz representation theorem. The inner product of any two elements $h_1, h_2 \in H$ is denoted by $\langle h_1, h_2 \rangle_H$ and the duality pairing of x^* and $x \in E$ is denoted by $\langle x^*, x \rangle_E$. All the vector spaces we shall consider are assumed to be real. We write $\mathcal{L}(H, E)$ to denote the space of bounded linear operators from H to E. The adjoint of the operators in $\mathcal{L}(H, E)$ are operators in $\mathcal{L}(E^*, H)$.

Given any measurable space $(S, \mathscr{S})$, recall that a mapping $S \to E$ is said to be simple if it is measurable and takes finitely many values. Further, we shall say that it is strongly measurable if it is the pointwise limit of a sequence of simple functions. If $(S, \mathscr{S}, \lambda)$ is a σ-finite measure space, for a given $p \in (1, +\infty)$ we write $L^p(S; E)$ to define the linear space of measurable functions $f : S \to E$ such that $\int_S \|f\|_E^p d\lambda < +\infty$, by identifying those functions that equal λ-almost everywhere. Endowed with the norm

$$\|f\|_{L^p(S;E)} \triangleq \left\{ \int_S \|f\|_E^p d\lambda \right\}^{1/p}, \qquad \text{for any } f \in L^p(S; E),$$

the space $L^p(S; E)$ is a Banach space.

Throughout, we fix $(\Omega, \mathscr{F}, \mathbb{P})$ to be a reference complete probability space. For any $X \in L^1(\Omega; E)$ we write $\mathbb{E}\{X\}$ to denote its expected value with respect to $\mathbb{P}$.

5.1 UMD SPACES AND RADONIFYING OPERATORS

In order to construct a theory of stochastic integration within a Banach space framework, a natural tool is provided by the spaces of the so-called γ-radonifying operators. This class of operators has been firstly characterized in the works by Gel'Fand [89], Segal [161] and Gross [96]. Besides, this theory is sound in the particular case when dealing with a specific class of spaces in which martingale differences are unconditional, nowadays known as UMD Banach spaces. This class of spaces was initially characterized in the works of Burkholder [32] and Bourgain [23]. We refer to Burkholder's review article [33] for an overview of the theory of UMD Banach spaces and its applications in stochastic and harmonic analysis.

Throughout this section we briefly discuss the characterization of the UMD Banach spaces and we review the main results that arise when considering the class of γ-radonifying operators.

UMD spaces and martingale type. We call martingale difference sequence in E any sequence $d_1, d_2, ...$ that is a difference sequence of a E-valued martingale, i.e. such that $d_n = M_n - M_{n-1}$ for $n \in \mathbb{N}$, where $M_0, M_1, ...$ is some E-valued martingale with $M_0 = 0$ a.s. Further, the sequence $d_1, d_2, ...$ is said to be a L^p-martingale difference sequence if the it is a difference sequence of a E-valued L^p-martingale.

Proposition 5.1. *Every L^2-martingale difference sequence in H is orthogonal in $L^2(\Omega; H)$.*

Proof. See, e.g. Proposition 12.2. in [171]. □

We call signs sequence any sequence $\varepsilon_1, \varepsilon_2, ...$ such that $\varepsilon_n = \pm 1$, for any $n \in \mathbb{N}$. Martingale sequences provide a natural generalization of orthogal sequences. Indeed, if $d_1, d_2, ...$ is a L^2-martingale sequence in a Hilbert space H and $\varepsilon_1, \varepsilon_2, ...$ is a signs sequence, Proposition 5.1 gives

$$\mathbb{E}\left\| \sum_{i=1}^{n} \varepsilon_i d_i \right\|_H^2 = \mathbb{E}\left\| \sum_{i=1}^{n} d_i \right\|_H^2, \tag{5.1}$$

for any $n \in \mathbb{N}$. When dealing with Banach spaces, this property is generalized by the following definition.

Definition 5.1. A Banach space E is said to be a *UMD Banach space* if for some $p \in (1, +\infty)$ (equivalently, for all $p \in (1, +\infty)$) there exists a constant $\beta \geq 0$ such that for any L^p-martingale difference sequence $d_1, d_2, ...$ in E and any sign sequence $\varepsilon_1, \varepsilon_2, ...$ one has

$$\mathbb{E}\left\| \sum_{i=1}^{n} \varepsilon_i d_i \right\|_E^p \leq \beta^p \mathbb{E}\left\| \sum_{i=1}^{n} d_i \right\|_E^p, \tag{5.2}$$

for any $n \in \mathbb{N}$.

The term "UMD" stands for "unconditional martingale differences". It is worth to be highlighted that as a direct consequence of Proposition 5.1 every Hilbert space is a UMD space.

We call Rademacher sequence any sequence $r_1, r_2, ...$ of independent random variables taking values ± 1 with probability $1/2$.

Definition 5.2. Let $1 \leq p \leq 2$. A Banach space E has *type* p if there exists a constant δ such that

$$\mathbb{E}\left\| \sum_{i=1}^{n} r_i x_i \right\|_E^p \leq \beta^p \sum_{i=1}^{n} \|x_i\|_E^p \tag{5.3}$$

for any $n \in \mathbb{N}$ and any finite sequence $x_1, x_2, ..., x_n$ in E.

The last possible constant β in (5.3) is said to be the *type* p *constant* of E and it will be denoted by β_p. It is worth to be highlighted that every Banach space has martingale type 1 and moreover any Hilbert space has martingale type 2.

Radonifying Operators. For any $h \in H$ and $x \in E$, we write $h \otimes x$ to denote the operator in $\mathcal{L}(H, E)$ defined by

$$(h \otimes x)k \triangleq x\langle h, k\rangle_H, \qquad \text{for any } k \in H.$$

Let $H \otimes E$ denote the linear space of finite rank operators from H to E, i.e. the operators of the form $\sum_{i=1}^{n} h_i \otimes x_i$, where $x_1, ..., x_n \in E$ and the vectors $h_1, ..., h_n \in H$ are orthonormal.

Let $\gamma_1, \gamma_2, ...$ be a sequence of real-valued independent standard Gaussian random variables and define the norm on $H \otimes E$ given by

$$\left\| \sum_{i=1}^{n} h_i \otimes x_i \right\|_{\gamma(H,E)} \triangleq \left\{ \mathbb{E} \left\| \sum_{i=1}^{n} \gamma_i x_i \right\|_{\gamma(H,E)}^2 \right\}^{1/2} \tag{5.4}$$

for any finite rank operator $\sum_{i=1}^{n} h_i \otimes x_i$.

DEFINITION 5.3. We define the Banach space $\gamma(H, E)$ as the completion of $H \otimes E$ with respect to the norm (5.4).

The space $\gamma(H, E)$ is an operator ideal in $\mathcal{L}(H, E)$, i.e. the following ideal property holds true.

Lemma 5.1. *Let H_1 and H_2 be Hilbert spaces and E_1 and E_2 Banach spaces. Let $S \in \mathcal{L}(H_2, H_1)$ and $T \in \mathcal{L}(E_1, E_2)$. Then, if $\vartheta \in \gamma(H_1, E_1)$ one has $T\vartheta S \in \gamma(H_2, E_2)$ and moreover the following norm equality holds true,*

$$\|T\vartheta S\|_{\gamma(H_2,E_2)} \leq \|T\|_{\mathcal{L}(E_1,E_2)} \|\vartheta\|_{\gamma(H_1,E_1)} \|S\|_{\mathcal{L}(H_2,H_1)}.$$

Proof. See, e.g., Theorem 6.2 in [172]. □

In the case when $\vartheta \in \gamma(H, E)$ and $x^* \in E^*$, we write $\langle x^*, \vartheta \rangle_E$ to denote the H-valued dual pairing between $\gamma(H, E)$ and E^*, defined by setting $\langle x^*, \vartheta \rangle_E \triangleq \vartheta^* x^*$, where ϑ^* denotes the Banach space adjoint operator of ϑ. On the other hand, from Lemma 5.1, one has that

$$\|\langle x^*, \vartheta \rangle_E\|_H \leq \|x^*\|_{E^*} \|\vartheta\|_{\gamma(H,E)}. \tag{5.5}$$

We recall that, if E is further assumed to be a Hilbert space, then $\gamma(H, E)$ boils down to the space of Hilbert-Schmidt operators mapping H into E. Hence, we have the natural identifications $\gamma(H, \mathbb{R}) = H$ and $\gamma(\mathbb{R}, E) = E$, endowed with the related norms.

Proposition 5.2. *Let H be a separable Hilbert space and fix $1 \leq p < +\infty$. Then, the mapping*

$$F \in L^p(\Omega; \gamma(H, E)) \mapsto (h \in H \mapsto F(\cdot)h),$$

defines the Banach spaces isomorphism

$$L^p(\Omega;\gamma(H,E)) \simeq \gamma(H;L^p(\Omega,E)). \tag{5.6}$$

Proof. See, e.g., Proposition 2.6 in [174]. □

Wiener space. Throughout, we show the connection between the class of γ-radonifying operators and the characterization of the Gaussian measures on E induced by a cylindrical Gaussian measure on H, as obtained by Gross [96] in terms of the so-called measurable semi-norms on H. The cited work marked the beginning of the study of the Gaussian measures on Banach spaces.

Let $\mathscr{R}_H$ be the ring of the cylindrical sets in H, i.e. the collection of the measurable sets of the form $T^{-1}(B)$, for some $T \in \mathcal{L}(H,\mathbb{R}^n)$, varying $B \in \mathscr{B}(\mathbb{R}^n)$ and $n \geq 1$. In particular, for any $T \in \mathcal{L}(H,\mathbb{R}^n)$, we define $\langle\cdot,\cdot\rangle_T$ to be the inner product on $\mathbb{R}^n$ with respect to which the operator T is indeed an isometry from $(\operatorname{Ker} T)^\perp$ to $\mathbb{R}^n$. Moreover, for any fixed $n \geq 1$, let $\mathscr{B}_T(\mathbb{R}^n)$ be the Borel σ-algebra consistent with $\langle\cdot,\cdot\rangle_T$. We write $\gamma^{(n)}$ to denote the standard Gaussian measure on the space $(\mathbb{R}^n, \mathscr{B}_T(\mathbb{R}^n))$. The *canonical cylindrical Gaussian measure* γ_H associated to the Hilbert space H is defined as the application

$$\gamma_H(T^{-1}(B)) \triangleq \gamma^{(n)}(B), \qquad \text{for any } B \in \mathscr{B}_T(\mathbb{R}^n). \tag{5.7}$$

It is well known fact that the application γ_H generally fails to be countably additive on $\mathscr{R}_H$ when the space H is infinite-dimensional, (cf. [160], Corollary of Lemma 6). Besides, L. Gross [96] proposed to define measure starting from γ_H, by generalizing the Wiener's original construction of the standard Wiener measure on the space of the paths. Let $\|\cdot\|_H$ be a continuous norm on H and define E to be the Banach space obtained as the completion of H with respect to $\|\cdot\|_H$.

As a direct consequence, there exists a continuous linear injection $\imath : H \to E$ such that the projection $\imath(\gamma_H) \triangleq \gamma_H \circ \imath^{-1}$ defines a positive definite function on the ring $\mathscr{R}_E$ of cylindrical sets in E. Let $\sigma(\mathscr{R}_E)$ be the σ-algebra generated by the ring $\mathscr{R}_E$ and $\mathscr{E}$

the Borel σ-algebra generated by $\|\cdot\|_H$ on E in the common way. Moreover, note that $\mathscr{E}$ and $\sigma(\mathscr{R}_E)$ coincide, since the space E is assumed to be separable. The norm $\|\cdot\|_H$ is called *admissible* if the application $\imath(\gamma_H)$ extends to a Borel measure γ on $(E, \mathscr{E})$. Gross proved that $\|\cdot\|_H$ is always admissible in the case when it is *measurable*, i.e. such that given any $\varepsilon > 0$, there exists a finite dimensional projection π_0 of H such that $\gamma_H\{h \in H : \|\pi h\| > \varepsilon\} < \varepsilon$, for any finite dimensional projection π orthogonal to π_0, (cf. [96], Theorem 1).

It follows directly from (5.7) that the measure γ obtained in such a way is Gaussian and for any $x^* \in E^*$, the projection $x^*(\gamma) \triangleq \gamma \circ (e^*)^{-1}$ provides a centered Gaussian measure on the real line with variance given by $\|x^*\|^2_{E^*}$. The following result gives that the linear injection $\imath$ belongs to the space $\gamma(H, E)$.

Proposition 5.3. *Let $\vartheta : H \to E$ be a linear operator and γ_H the canonical Gaussian cylindrical measure on H. Then, $\vartheta \in \gamma(H, E)$ if and only if $\gamma_H \circ \vartheta^{-1}$ extends to a Gaussian measure on $(E, \mathscr{E})$.*

Proof. See, e.g., Theorem 5.2 jointly with Definition 5.1 in [155]. □

The triple $(E, H, \imath)$ is usually called *abstract Wiener space*, and H is said the *Cameron-Martin* space associated to the measure γ. Moreover, every separable Banach space E endowed with a Gaussian measure γ always contains a densely embedded Hilbert space H such that the triple (E, H, γ) forms an abstract Wiener space, (cf. [160], Theorem 2).

Proposition 5.4. *If H has infinite dimension, then $\gamma(H) = 0$.*

Proof. See, e.g., Theorem 1.3 in [12]. □

Let $Q : E^* \to E$ be the linear operator defined by

$$x^* \mapsto Qx^* \triangleq \int_E x\langle x^*, x\rangle \gamma(dx), \qquad \text{for any } x^* \in E^*, \tag{5.8}$$

where the latter integral has to be regarded in the Bochner sense, since the integrand turns out to be E-valued. Note that Q is positive and symmetric, i.e. such that

$\langle x^*, Qx^* \rangle \geq 0$ and $\langle x^*, Qy^* \rangle = \langle Qx^*, y^* \rangle$, for any x^* and y^* in E^*. The operator Q is said to be the *covariance operator* of γ.

Define $H(\gamma)$ to be the Hilbert space given by the completion of the range $R(E^*)$ with respect to the inner product

$$\langle Qx_1^*, Qx_2^* \rangle_{H(\gamma)} \triangleq \langle x_1^*, Qx_2^* \rangle \qquad \text{for any } x_1^*, x_2^* \in E^*. \tag{5.9}$$

The space $H(\gamma)$ is usually called the *reproducing kernel Hilbert space* generated by γ. Note that $H(\gamma)$ is separable since the space E is assumed to separable. Further, the linear inclusion mapping from the range $Q(E^*)$ to E is easily seen to be continuous with respect to the inner product (5.9) and the norm in E respectively. Hence, it can be extended uniquely to a bounded linear injection $\imath_Q : H(\gamma) \to E$. On the other hand, since the support of the measure γ coincides with the entire space E, the range of $\imath_Q$ is dense in E. Thus, the operator Q enjoys the decomposition $Q = \imath_Q^* \circ \imath_Q$, where the adjoint $\imath_Q^*$ is also one-to-one with dense range, (cf. [155], §4), i.e.

$$E^* \xrightarrow{\imath_R^*} H(\gamma)^* \cong H(\gamma) \xrightarrow{\imath_R} E.$$

Proposition 5.5. *Let $\vartheta : H \to E$ be a linear and bounded operator. Then, $\vartheta \in \gamma(H, E)$ if and only if $\theta \circ \theta^*$ is the covariance operator associate to some Gaussian measure on E.*

Proof. See, e.g., Proposition 1.1 in [169]. □

5.2 MALLIAVIN CALCULUS

The theory of Malliavin calculus may be traced back to the pioneering work developed by Malliavin [130], who provided a probabilistic proof of the Hörmander theorem. The Malliavin calculus naturally generalizes to the Hilbert space framework. We refer to [40] for an exhaustive accounting applied to the theory of stochastic differential equation in the infinite dimensional setup. In the framework of Hilbert spaces, details of this matter may be also found in [80, 95, 121] and the references therein.

Throughout this section, we discuss the theory of Malliavin calculus and stochastic integration that has been proposed in the recent years when dealing in a Banach space setup. In this respect, we manly refer to the works by Mass and van Neerven [128]. Other references are the work of Pronk and Varaar [149] and Maas [127].

Isonormal processes. Recall that an E-valued random variable X is said to be Gaussian if the real valued variable $\langle x^*, X\rangle_E$ is Gaussian, for any $x^* \in E^*$. Throughout, any Gaussian random variable X is assumed to be cantered, i.e. $\mathbb{E}\{X\} = 0$.

Definition 5.4 (Isonormal Gaussian Process). We call *H-isornomal Gaussian process* on the space $(\Omega, \mathscr{F}, \mathbb{P})$ any mapping $\mathscr{W} : H \to L^2(\Omega)$ that satisfies the following properties

(i) the random variable $\mathscr{W}(h)$ is Gaussian, for any $h \in \mathscr{H}$;

(ii) for any $h, k \in H$ one has $\mathbb{E}\{\mathscr{W}(h)\mathscr{W}(k)\} = \langle h, k\rangle_H$.

Note that from (ii) in Definition 5.4, one has

$$\mathbb{E}\{|\mathscr{W}(c_1 h_1 + c_2 h_2) - (c_1 \mathscr{W}(h_1) + c_2 \mathscr{W}(h_2))|^2\} = 0, \qquad \text{for any scalar } c_1, c_2 \in \mathbb{R} \text{ and } h_1, h_2 \in H,$$

which implies that the process $\mathscr{W}$ is linear. As a consequence, for any $h_1, ..., h_n \in H$, the random variables $\mathscr{W}(h_1), ..., \mathscr{W}(h_n)$ turn out to be jointly Gaussian and hence independent if and only if $h_1, ..., h_n \in H$ are orthogonal.

Remark. According to Kolmogorov's extension theorem, once a separable Hilbert space H is given, one can always construct a probability space and a H-inonormal Gaussian process. Indeed, let $h_1, h_2, ...$ be an orthogonormal system in H and $\eta_1, \eta_2, ...$ a sequence of independent standard Gaussian real-valued random variables. For any $h \in H$, set $\mathscr{W}(h) \triangleq \sum_{n\geq 1} \langle h, h_n\rangle_H \eta_n$, where the series converges $\mathbb{P}$-a.s. and in $L^2(\Omega)$, since $\sum_{n\geq 1} \langle h, h_n\rangle < \infty$. The process $\mathscr{W}$ verifies the properties in Definition 5.4.

Given any Banach space E, any isonormal Gaussian process $\mathscr{W} : H \to L^2(\Omega)$ naturally induces a linear mapping $\tilde{\mathscr{W}}$ from $H \otimes E$ to $L^2(\Omega) \otimes E$, by setting

$$\tilde{\mathscr{W}}(h \otimes x) \triangleq \mathscr{W}(h) \otimes x, \qquad \text{for any } h \in H \text{ and } x \in E.$$

This is made clear by the following result.

Proposition 5.6. *Any H-isonormal process $\mathscr{W} : H \to L^2(\Omega)$ induces an isometry $\tilde{\mathscr{W}} : \gamma(H, E) \to L^2(\Omega; E)$.*

Proof. See, e.g., Proposition 3.9 in [172]. □

The following result provides an useful mapping property of the space $\gamma(H, E)$.

Proposition 5.7. *Every $x^* \in E^*$ extends to a bounded operator $\tilde{x}^* : \gamma(H, E) \to H$, by setting*

$$\tilde{x}^*(h \otimes x) \triangleq \langle x^*, x\rangle_H h, \qquad \textit{for any } h \otimes x \in H \otimes E. \tag{5.10}$$

Moreover, the following diagram commutes

$$\begin{tikzcd} \gamma(H,E) \arrow[r, "\tilde{\mathscr{W}}"] \arrow[d, "\tilde{x}^*"'] & L^2(\Omega; E) \arrow[d, "x^*"] \\ H \arrow[r, "\mathscr{W}"] & L^2(\Omega) \end{tikzcd}$$

Proof. See, e.g., Proposition 3.12 in [172]. □

The Malliavin derivative. Throughout, we fix a H-isonormal process $\mathscr{W}$. We write $\mathscr{C}_b^\infty(\mathbb{R}^n)$ to denote the the vector space of all bounded and real-valued functions defined on $\mathbb{R}^n$ that have bounded derivative of any order. We call E-valued *smooth random variable* any function $X : \Omega \to E$ if the following identity holds

$$X = f(\mathscr{W}(h_1), ..., \mathscr{W}(h_n)) \otimes x, \tag{5.11}$$

for some $f \in \mathscr{C}_b^\infty(\mathbb{R}^n)$ and $h_1, ..., h_n \in \mathscr{H}$. Furthermore, we say that a smooth random variable X of the form (5.11) is *compactly supported* if the the function f is compactly supported. The class of smooth E-valued random variables and compactly supported smooth E-valued random variables will be denoted by $\mathscr{S}(\Omega) \otimes E$ and $\mathscr{S}_c(\Omega) \otimes E$, respectively.

Lemma 5.2. *$\mathscr{S}_c(\Omega) \otimes E$ is dense in $L^p(\Omega; E)$, for $1 \leq p < \infty$.*

Proof. See, e.g., Lemma 3.1. in [128]. □

Let X be a E-valued smooth random variable of the form (5.11). Then, the *Malliavin derivative* of X is defined as the $\gamma(H, E)$-valued random variable DX given by

$$DX \triangleq \sum_{i=1}^{n} \partial_i \psi(\mathscr{W}(h_1), ..., \mathscr{W}(h_n)) \otimes (h_i \otimes x). \tag{5.12}$$

Note that this definition naturally extends to the entire space $\mathscr{S}(\Omega) \otimes E$, by linearity. Besides, it is worth to be highlighted that the Malliavin derivative (5.12) may be regarded as an E-valued stochastic process $D_h X$, for $h \in \mathscr{H}$, by setting $D_h X \triangleq (DX)h$.

The following result plays a central role.

Proposition 5.8. *Fix $p \in (1, +\infty)$. The Malliavin derivative D is a closable operator from $L^p(\Omega; E)$ to $L^p(\Omega; \gamma(H, E))$.*

Proof. See, e.g., Proposition 3.3. in [128]. □

Here and in the sequel, we still denote by D the closure of the Malliavin derivative D and we write $\mathbb{H}^{1,p}(E)$ to denote its domain in $L^p(\Omega; E)$, which is a Banach space when endowed with the graph norm

$$\|X\|_{\mathbb{H}^{1,p}(E)} \triangleq \{\|X\|^p_{L^p(\Omega;E)} + \|DX\|^p_{L^p(\Omega;\gamma(H,E))}\}^{1/p}, \qquad \text{for any } X \in \mathbb{H}^{1,p}(E). \tag{5.13}$$

In this respect the domain $\mathbb{H}^{1,p}(E)$ is nothing but the completion of the class of the smooth E-valued random variables in $L^p(\Omega; F)$ with respect to the norm (5.14). Besides, it is worth highlighting that the topology induced by the norm (5.14) is finer than the standard topology associated to $L^p(\Omega; E)$; indeed, the two topologies would be equivalent only if the Malliavin derivative D were bounded on $L^p(\Omega; E)$, which is not the case in general.

Moreover, for any $k \geq 2$, we denote by $D^k \triangleq D \circ ... \circ D$ the k-fold composition of the Malliavin derivative operator $D : \mathbb{H}^{1,p}(E) \to L^p(\Omega, \gamma(H, E))$. The operator D^k takes values in the space $L^p(\Omega; \gamma^k(H, E))$, where we set $\gamma^1(H, E) \triangleq \gamma(H, E)$ and $\gamma^{n+1}(H, E) \triangleq \gamma(H, \gamma^n(H, E))$ by recursion for any $n \in \mathbb{N}$. The domain $\mathbb{H}^{k,p}(E)$ of D^k coincides with the closure of the class of smooth E-valued random variables in $L^p(\Omega; E)$ with respect to the norm

$$\|X\|_{\mathbb{H}^{1,p}(E)} \triangleq \left\{ \|X\|^p_{L^p(\Omega;E)} + \sum_{n=1}^{k} \|D^n X\|^p_{L^p(\Omega;\gamma^n(H,E))} \right\}^{1/p}. \tag{5.14}$$

The following result provides a chain rule for the Malliavin derivative operator.

Proposition 5.9. *Let E_1 and E_2 be Banach spaces. Assume E_2 to be a UMD space and fix $p \in (1, +\infty)$. Given a Fréchet differentiable function $\varphi : E_1 \to E_2$ such that its Frechét derivative $\nabla\varphi : E \to \mathcal{L}(E_1, E_2)$ is continuous and bounded, if $X \in \mathbb{H}^{1,p}(E_1)$ then $\varphi(X) \in \mathbb{H}^{1,p}(E_2)$. Moreover, one has*

$$D\varphi(X) = \nabla\varphi(X)DX. \tag{5.15}$$

Proof. See, e.g., Proposition 3.8 in [149]. □

Moreover, the following product rule will be useful later on

Proposition 5.10. *Fix $p \in (1, +\infty)$ and let $q, r \in (1, +\infty)$ be such that $p^{-1} + q^{-1} = r^{-1}$. If $X \in \mathbb{H}^{1,p}(E)$ and $Y \in \mathbb{H}^{1,q}(E^*)$ then $\langle X^*, X\rangle_E \in \mathbb{H}^{1,r}(\mathbb{R})$ and*

$$D\langle X^*, X\rangle_E = \langle DX^*, X\rangle_E + \langle X^*, DX\rangle_E. \tag{5.16}$$

Proof. See, e.g., Lemma 3.6 in [149]. □

The divergence operator. For any fixed $p \in (1, +\infty)$, the divergence operator δ is defined as the part of the adjoint operator D^* in $L^p(\Omega; \gamma(H, E))$ which maps into

$L^p(\Omega; E)$. To make this precise, let $q \in (1, +\infty)$ be the conjugate exponent of p, i.e. such that $p^{-1} + q^{-1} = 1$. Assume D to be the Malliavin derivative on $L^q(\Omega; E^*)$, that is densely defined as a $L^q(\Omega; \gamma(H, E^*))$-valued closed operator with domain $\mathbb{H}^{1,q}(E^*)$. Thus, let $\mathrm{dom}_p(\delta)$ be the set of $X \in L^p(\Omega; \gamma(H, E))$ for which there exists $Z \in L^p(\Omega; E)$ such that

$$\mathbb{E}\langle X, DY\rangle_E = \mathbb{E}\langle Z, Y\rangle_E, \qquad \text{for any } Y \in \mathbb{H}^{1,p}(E^*). \tag{5.17}$$

The variable Z, when it exists, is uniquely defined and hence we set

$$\delta(X) \triangleq Z, \qquad \text{for any } X \in \mathrm{dom}_p(\delta).$$

The divergence operator δ is easily seen to be closed and densely defined (cf. Lemma 4.1 in [128]).

Proposition 5.11. *Fix $p \in (1, +\infty)$ and assume E to be a UMD space. The operator δ is continuous from $\mathbb{H}^{1,p}(\gamma(H, E))$ to $L^p(\Omega; E)$.*

Proof. See, e.g., Proposition 6.10 in [127]. □

The following result provides a useful commutation relation between the operators D and δ.

Proposition 5.12. *Fix $p \in (1, +\infty)$ and assume E to be a UMD space. If $X \in \mathbb{H}^{2,p}(\gamma(H, E))$, then $\delta(X) \in \mathbb{H}^{1,p}(E)$. Moreover,*

$$D\delta(X) = X + \delta(DX), \qquad \textit{for any } X \in \mathbb{H}^{2,p}(\gamma(H, E)). \tag{5.18}$$

Proof. See, e.g., Proposition 4.4 in in [149]. □

5.3 STOCHASTIC INTEGRATION

In this Section we briefly present the theory of stochastic integration in UMD Banach spaces. We mainly refer to the works of van Neerven, Veraar and Weis [174, 177] and

[178]. In the sequel, we let I denote the unit interval on the real line and thus we write $\mathscr{H} = L^2(I; H)$.

Measurablility and representation. Throughout, an E-valued process is a one-parameter family of E-valued random variables indexed by I. We shall say that a process is strongly measurable if it consists in a one-parameter family of strongly measurable random variables. In most cases, we identify the generic E-valued process with the induced map $\Omega \times I \to E$.

Any $\mathcal{L}(H,E)$-valued process $X \triangleq \{X_t : t \in I\}$ is said to be H-strongly measurable if the E-valued process $Xh \triangleq \{X_t h, t \in I\}$ is strongly measurable, for any $h \in H$. Further, we shall say that an H-strongly measurable $\mathcal{L}(H,E)$-valued process X belongs to $\mathscr{H}$ scalarly a.s. if for any $x^* \in E^*$ the function $t \in I \mapsto X_t(\omega)^* x^* \in E$ belongs to $\mathscr{H}$ for almost any $\omega \in \Omega$. Such a process is said to represent an H-strongly measurable $\mathcal{L}(\mathscr{H}, E)$-valued variable Z in the special case when for any $x \in E^*$ and any $f \in \mathscr{H}$ one has

$$\langle x^*, Zf\rangle_E = \int_I \langle X_t^* x^*, f_t\rangle_H dt, \qquad \text{a.s.} \tag{5.19}$$

The notion of representation above is taken from [174].

Equivalently, an H-strongly measurable $\mathcal{L}(H,E)$-valued process X which belongs to $\mathscr{H}$ scalarly a.s. represents an H-strongly measurable $\mathcal{L}(\mathscr{H}, E)$-valued variable Z if the following identity holds (cf. [174], Lemma 2.7),

$$\langle x^*, Zf\rangle_E = \int_I \langle x^*, X_t f_t\rangle_E dt, \qquad \text{a.s.}, \tag{5.20}$$

for any $x^* \in E^*$ and any $f \in \mathscr{H}$. Thus, given any simple function $Y : I \to \gamma(H,E)$, we set

$$I_Y f \triangleq \int_I Y_t f_t dt, \qquad \text{for any } f \in \mathscr{H}, \tag{5.21}$$

where the right hand side in (5.21) is defined in an elementary way. The following embedding result via representation turns out to be convenient.

Lemma 5.3. *If E has type 2, the mapping $\imath_2 : Y \mapsto I_Y$ given by (5.21) admits an unique extension to a continuous embedding*

$$\imath_2 : L^2(I, \gamma(H, E)) \hookrightarrow \gamma(\mathscr{H}, E), \tag{5.22}$$

with operatorial norm satisfying $\|\imath_2\| \leq \beta_2$, where β_2 denotes the type 2 constant of E.

Proof. See, e.g., Lemma 6.1 in [178]. □

Stochastic integral. Later on, a cylindrical H-Wiener process on Ω is a mapping $W : H \to L^2(\Omega)$ satisfies the following condition.

(i) For any $h \in H$, the process $Wh \triangleq \{W_t h : t \in I\}$ is a standard Brownian motion;

(ii) For any $t, s \in I$ and $h_1, h_2 \in H$, one has that $\mathbb{E}\{W_t h_1 W_s h_2\} = (s \wedge t)\langle h_1, h_2\rangle_H$.

It is worth to be noted that if $\mathscr{W}$ is an $\mathscr{H}$-Wiener process, then the process $W : H \to L^2(\Omega)$ defined by the following identity

$$W_t h \triangleq \mathscr{W}(\mathbb{1}_{[0,t]} \otimes h), \qquad \text{for any } t \in I \text{ and } h \in H, \tag{5.23}$$

defines a cylindrical H-Wiener process.

We set $\mathscr{G}^W \triangleq \{\mathscr{G}_t^W : t \in I\}$ to be the augmented filtration generated by the H-Wiener process W, i.e. generated by the Brownian motions Wh, varying $h \in H$. Further, we say that an E-valued process is adapted when it is adapted to the filtration $\mathscr{G}^W$. On the other hand, an H-strongly measurable $\mathcal{L}(H, E)$-valued process X is said to be adapted if the E-valued process Xh is adapted, for any $h \in H$.

Following [174, 177] and [128], we say that a process $X : \Omega \times I \to \gamma(H, E)$ is an elementary adapted process with respect to the filtration $\mathscr{G}^W$ if it admits the following representation

$$X_t = \sum_{i=1}^{m}\sum_{j=1}^{n} \mathbb{1}_{(t_{i-1}, t_i]}(t)\mathbb{1}_{A_{ij}} \sum_{k=1}^{l} h_k \otimes x_{ijk}, \qquad \text{for any } t \in I, \tag{5.24}$$

where $0 \leq t_0 < \cdots < t_n \leq 1$, the sets $A_{ij} \in \mathscr{G}_{t_{i-1}}$ are disjoint for each j, and $h_1, ...h_k \in \mathscr{H}$ are orthonormal. Thus, any elementary adapted process defines an element of $L^p(\Omega; \gamma(\mathscr{H}, E))$, for $p \in [1, +\infty)$. We write $L^p_{\mathscr{G}^W}(\Omega; \gamma(\mathscr{H}, E))$ to denote their closure in $L^p(\Omega; \gamma(\mathscr{H}, E))$.

The stochastic integral of an elementary adapted process X of the form (5.24) with respect to W is defined as the $L^p(\Omega; E)$-valued variable $\mathcal{I}_W(X)$ defined by

$$\mathcal{I}_W(X) \triangleq \int_I X_t dW_t \triangleq \sum_{i=1}^{m} \sum_{j=1}^{n} \sum_{k=1}^{l} \mathbb{1}_{A_{ij}} (W_{t_i} h_k - W_{t_{i-1}} h_k) \otimes x_{ijk}. \tag{5.25}$$

Note that $\mathcal{I}_W(X)$ does not depend on the choice of the adapted elementary process X.

The following result provides an extension of the stochastic integral with respect to W to the entire space $L^p_{\mathscr{G}^W}(\Omega; \gamma(\mathscr{H}, E))$.

Theorem 5.1 (Itô isomorphism)**.** *Let E be a UMD space and fix $p \in (1, +\infty)$. The mapping $X \mapsto \mathcal{I}_W(X)$ defined by (5.25) admits a unique extension to a bounded operator*

$$\mathcal{I}_W : L^p_{\mathscr{G}^W}(\Omega; \gamma(\mathscr{H}, E)) \to L^p(\Omega; E). \tag{5.26}$$

Moreover, the operator (5.26) defines a isomorphism of Banach spaces.

Proof. See, e.g., Theorem 3.5 in [174]. □

The following theorem provides a characterization of the class of stochastically integrable processes.

Theorem 5.2. *Let E be a UMD space and fix $p \in (1, +\infty)$. Further, consider an H-strongly measurable and adapted $\mathcal{L}(H, E)$-valued process $X = \{X_t : t \in I\}$ that belongs to $L^p(I; H)$ scalarly. The following assertions are equivalent:*

i. *There exists a sequence $X_1, X_2, ...$ of elementary adapted processes such that*

   a. *for any $h \in H$ and $x^* \in E^*$ one has that $\lim_n \langle x^*, X_n h\rangle_E = \langle x^*, Xh\rangle_E$ in measure on $I \times \Omega$,*

b. *there exists a strongly measurable random variable* $Y \in L^p(\Omega; E)$ *satisfying the following identity*

$$Y = \lim_n \mathcal{I}_W(X_n), \qquad \textit{in } L^p(\Omega; E);$$

ii. *there exists a strongly measurable random variable* $Y \in L^P(\Omega; E)$ *such that for any* $x^* \in E^*$ *one has*

$$\langle x^*, Y \rangle_E = \mathcal{I}_W(X^* x^*), \qquad \textit{in } L^p(\Omega);$$

iii. *The process* X *represents an element* $Z \in L^p(\Omega; \gamma(\mathscr{H}, E))$.

Further, the random variables Y *in* i. *and* ii. *are uniquely determined and they coincide as elements of* $L^P(\Omega; E)$. *Moreover, the variable* Z *in point* iii. *belongs to the space* $L^p_{\mathscr{G}^W}(\Omega; \gamma(\mathscr{H}, E))$ *and one has that*

$$Y = \mathcal{I}_W(Z), \qquad \textit{in } L^p(\Omega; E). \tag{5.27}$$

Proof. See Theorem 3.6 in [174]. □

Any $\mathcal{L}(H, E)$-valued process X satisfying the equivalent conditions of Theorem 5.2 is said L^p-*stochastically integrable* and the random variable $Y = \mathcal{I}_W(Z)$ is called the *stochastic integral* of X with respect to W. Further, we shall use the notation

$$Y = \mathcal{I}_W(Z) \triangleq \int_I X_t dW_t. \tag{5.28}$$

Note that by Lemma 5.3 if the space E has type 2 then there exists an inclusion $L^2(I, \gamma(H, E)) \hookrightarrow \gamma(\mathscr{H}, E)$ and this inclusion is given by representation. Hence, when dealing with UMD Banach spaces with type 2, from Theorem 5.2 we obtain the following useful stochastic integrability characterization.

Corollary 5.1. *Let* E *be a UMD space with type* 2 *and fix* $p \in (1, +\infty)$. *Every* H*-strongly measurable and adapted process* X *that belongs to* $L^p(\Omega; L^2(I; \gamma(H, E)))$ *is* L^p*-stochastically integrable with respect to* W.

The following result provides a useful series expansion representation of the stochastic integral (5.28).

Proposition 5.13 (Series expansion)**.** *Let E be a UMD space and fix $p \in (1, +\infty)$. Fix an H-strongly measurable and adapted $\mathcal{L}(H, E)$-valued process X that is stochastically integrable with respect to W. Then, one has that for all $h \in H$ the process $Xh \triangleq \{X_t h : t \in I\}$ is stochastically integrable with respect to Wh. Further, given an orthonormal basis $h_1, h_2, \ldots$ in H, the following representation holds*

$$\mathcal{I}_W(X) = \sum_{n \geq 1} \mathcal{I}_{Wh_n}(Xh_n), \tag{5.29}$$

where the series converges unconditionally in $L^p(\Omega; E)$.

When considering the $\mathscr{H}$-isonormal process $\mathscr{W}$ that satisfies the identity (5.23), the divergence operator δ associated to the Mallivian derivative operator $D : L^p(\Omega; E) \to L^p(\Omega; \gamma(\mathscr{H}, E))$ turns out to be an extension of the stochastic integral $\mathcal{I}_W$. This is the content of the following theorem.

Theorem 5.3. *Let E be a UMD Banach space and fix $p \in (1, +\infty)$. The space $L^p_{\mathscr{G}^W}(\Omega; \gamma(\mathscr{H}, E))$ is contained in $dom_p(\delta)$ and*

$$\delta(X) = \mathcal{I}_W(X), \qquad \textit{for any } X \in L^p_{\mathscr{G}^W}(\Omega; \gamma(\mathscr{H}, E)).$$

Proof. See, e.g., Theorem 5.4 in [128]. □

A Clark-Ocone representation. For any step function $f : I \to \gamma(H, L^p(\Omega; E))$, the mapping $\pi^W_{\mathscr{G}}$ defined by the identity

$$(\pi_{\mathscr{G}^W} f)(t) \triangleq \mathbb{E}\{f(t) | \mathscr{G}^W_t\},$$

where $\mathbb{E}\{\cdot | \mathscr{G}^W_t\}$ is regarded as a bounded operator defined on $\gamma(H, L^p(\Omega; E))$. The mapping $\pi_{\mathscr{G}^W}$ admits a unique extension to a bounded operator defined on $\gamma(\mathscr{H}, L^p(\Omega; E))$ and such an extension defines a projection onto the subspace $L^p_{\mathscr{G}^W}(\Omega; \gamma(\mathscr{H}, E))$, (cf. see Lemma 6.5 in [128]). Throughout, with a slight abuse of notation, we shall denote the extension of $\pi_{\mathscr{G}^W}$ again by $\pi_{\mathscr{G}^W}$.

Theorem 5.4 (Clark-Ocone representation formula)**.** *Let E be a UMD space and fix $p \in (1, +\infty)$. The operator $\pi_{\mathscr{G}^W} \circ D$ admits a unique extension to a bounded and linear operator form $L^p(\Omega, \mathscr{G}_1^W; E)$ to $L^p_{\mathscr{G}^W}(\Omega; \gamma(\mathscr{H}, E))$. Moreover, the following identity holds*

$$X = \mathbb{E}\{X\} + \mathcal{I}_W((\pi_{\mathscr{G}^\mathscr{W}} \circ D)X), \qquad \text{for any } X \in L^P(\Omega, \mathscr{G}_1^W; E).$$

Further, the variable $(\pi_{\mathscr{G}^\mathscr{W}} \circ D)X$ is the unique element $Y \in L^p_{\mathscr{G}^W}(\Omega; \gamma(\mathscr{H}, E))$ satisfying $X = \mathbb{E}\{X\} + \mathcal{I}_W(Y)$.

Proof. See, e.g., Theorem 6.7 in [128]. □

5.4 STOCHASTIC EVOLUTION

In this section we present some useful results regarding the E-valued diffusion process defined in terms of the stochastic integral with respect to a cylindrical Wiener process. In this respect, in what follows we always assume E to be a UMD space with type 2. Further, we fix a separable Hilbert space H and an H-cylindrical Wiener process W toghether with the augmentation $\mathscr{G}^W$ of the filtration that is generates. We write $h_1, h_2, \ldots$ to denote the generic orthonormal basis in H. Furthermore, we still write $\mathscr{H} = L^2(I; H)$. In the sequel, a process X is adapted if it is adapted to the filtration $\mathscr{G}^W$. Moreover, we say X to be stochastically integrable if it is so with respect to W, as discussed in the previous section.

Although most of the results we present in this section are known, for convenience, we prove those ones we could not find in a suitable form in the existing literature.

Stochastic Evolution. Let $\xi_0 \in L^2(\Omega; E)$ be a strongly $\mathscr{G}_0^W$-measurable random variable. Consider an adapted and strongly measurable E-valued stochastic process $b = \{b_t : t \in I\}$, that belongs to $L^2(\Omega; L^2(I; E))$. Let $\sigma = \{\sigma_t : t \in I\}$ be some adapted and H-strongly measurable $\mathcal{L}(H, E)$-valued process that belongs to $L^2(\Omega; L^2(I; \gamma(H, E)))$. Moreover, we assume that the following conditions hold true,

$$\xi_0 \in \mathbb{H}^{1,2}(E), \quad b \in \mathbb{H}^{1,2}(L^2(I; E)), \quad \sigma \in \mathbb{H}^{2,2}(L^2(I; \gamma(H, E))). \tag{5.30}$$

The result below will be useful later on.

Lemma 5.4. *The process σ is well defined as an element of $L^2(\Omega;\gamma(\mathscr{H},E))$. Moreover, for $k=1,2$, we have that $\sigma\in\mathbb{H}^{k,2}(\gamma(\mathscr{H},E))$, and the following norm inequality holds,*

$$\|\sigma\|_{\mathbb{H}^{k,2}(\gamma(\mathscr{H},E))}\lesssim_E\|\sigma\|_{\mathbb{H}^{k,2}(L^2(I,\gamma(H,E)))}. \tag{5.31}$$

Proof. Since the space E is assumed to have type 2, Lemma 5.3 implies that there exists a continuous and linear embedding

$$\imath_2 : L^2(I,\gamma(H,E))\hookrightarrow\gamma(\mathscr{H},E),$$

with operatorial norm satisfying $\|\imath_2\|\le\beta_2$, where β_2 denotes the type 2 constant of E. Thus, the process σ turns out to be well defined as an element of $L^2(\Omega;\gamma(\mathscr{H},E))$.

Moreover, one has $\|\cdot\|_{\gamma(\mathscr{H},E)}\lesssim_E\|\cdot\|_{L^2(I;\gamma(H,E))}$, and hence for $k=1,2$, the following inequality holds true,

$$\|\cdot\|_{\mathbb{H}^{k,2}(\gamma(\mathscr{H},E))}\lesssim_E\|\cdot\|_{\mathbb{H}^{k,2}(L^2(I,\gamma(H,E)))}. \tag{5.32}$$

Thus, since we assumed that $\sigma\in\mathbb{H}^{2,2}(L^2(I;\gamma(H,E)))$, from inequality (5.32) we obtain that $\sigma\in\mathbb{H}^{1,2}(\gamma(\mathscr{H},E))$. $\square$

For any stochastically integrable $\mathcal{L}(H,\mathbb{R})$-valued process, the version of the Itô's isometry is regarded as follows.

Lemma 5.5. *Let $\psi\triangleq\{\psi_t : t\in I\}$ be an adapted H-strongly measurable and stochastically integrable process, taking values in $\mathcal{L}(H,\mathbb{R})$. Then, for any $t\in I$, one has*

$$\mathbb{E}\left\{\left|\int_0^t\psi_s dW_s\right|^2\right\}=\mathbb{E}\left\{\int_0^t\|\psi_s\|_H^2 ds\right\}.$$

Proof. First, notice that Proposition 5.13 implies that, for any $n\ge1$, the process $\psi h_n\triangleq\{\langle\psi_t,h_n\rangle_H : t\in I\}$ is stochastically integrable with respect to $Wh_n\triangleq\{W_t h_n : t\in I\}$, and for any $t\in I$ the following representation holds,

$$\int_0^t\psi_s dW_s=\sum_{n\ge1}\int_0^t\langle\psi_s,h_n\rangle_H dW_s h_n, \tag{5.33}$$

where the convergence of the series in (5.33) is understood in the topology of $L^2(\Omega)$.

Besides, since the processes Wh_n and Wh_m are independent for any $n, m \geq 1$ such that $n \neq m$, jointly with the Itô's isometry, for any $t \in I$ we have that,

$$\begin{aligned} \mathbb{E}\bigg\{ \int_0^t \langle \psi_s, h_n\rangle_H dW_s h_n \cdot \int_0^t \langle \psi_s, h_m\rangle_H dW_s h_m \bigg\} &= \delta_{nm}\mathbb{E}\bigg\{ \bigg| \int_0^t \langle \psi_s, h_n\rangle_H dW_s h_n \bigg|^2 \bigg\} \\ &= \delta_{nm}\mathbb{E}\bigg\{ \int_0^t |\langle \psi_s, h_n\rangle_H|^2 ds \bigg\}, \end{aligned} \tag{5.34}$$

where δ_{nm} denotes the Kronecker delta $\delta_{nm} = 1$ if $n = m$ and $\delta_{nm} = 0$ otherwise.

Thus, we have

$$\begin{aligned} \mathbb{E}\bigg\{ \bigg| \int_0^t \psi_s dW_s \bigg|^2 \bigg\} &\overset{\text{(i)}}{=} \mathbb{E}\bigg\{ \bigg| \sum_{n\geq 1} \int_0^t \langle \psi_s, h_n\rangle_H dW_s h_n \bigg|^2 \bigg\} \\ &= \sum_{n\geq 1}\sum_{m\geq 1} \mathbb{E}\bigg\{ \int_0^t \langle \psi_s, h_n\rangle_H dW_s h_n \cdot \int_0^t \langle \psi_s, h_m\rangle_H dW_s h_m \bigg\} \\ &\overset{\text{(ii)}}{=} \sum_{n\geq 1} \mathbb{E}\bigg\{ \int_0^t |\langle \psi_s, h_n\rangle_H|^2 ds \bigg\} \\ &= \mathbb{E}\bigg\{ \int_0^t \|\psi_s\|_H^2 ds \bigg\}, \end{aligned}$$

where in the step (i) we have used the representation (5.33) and in the step (ii) the identity (5.34). □

Lemma 5.6. *The process σ is stochastically integrable.*

Proof. Since $\sigma \in L^2(\Omega; L^2(I; \gamma(H, E)))$, then the result follows directly from Corollary 5.1. □

We consider the E-valued stochastic process $\xi \triangleq \{\xi_t : t \in I\}$ defined by,

$$\xi_t \triangleq \xi_0 + \int_0^t b_s ds + \int_0^t \sigma_s dW_s, \qquad \text{for any } t \in I. \tag{5.35}$$

The following result is taken from [149]; we include a proof for convenience.

Lemma 5.7. *For any $t \in I$, we have that $\xi_t \in \mathbb{H}^{1,2}(E)$.*

First, we prove the following Lemma.

Lemma 5.8. *For any $t \in I$, $\xi_t \in L^2(\Omega; E)$ is well defined. Moreover, we have*

$$\sup_{t\in I} \|\xi_t\|^2_{L^2(\Omega;E)} < \infty.$$

Proof of Lemma 5.8. Since the operator δ is linear and continuous from $\mathbb{H}^{1,2}(\gamma(\mathscr{H}, E))$ to $L^2(\Omega; E)$ due to Proposition 5.11, then by Lemma 5.4 that

$$\|\delta(\sigma)\|_{L^2(\Omega;E)} \lesssim_E \|\sigma\|_{\mathbb{H}^{1,2}(\gamma(\mathscr{H},E))} \lesssim_E \|\sigma\|_{\mathbb{H}^{1,2}(L^2(I;\gamma(H,E)))}.$$

As a consequence, we obtain

$$\begin{aligned}
\sup_{t\in I} \|\xi_t\|^2_{L^2(\Omega;E)} &\leq \|\xi_0\|^2_{L^2(\Omega;E)} + \|b\|^2_{L^2(\Omega;L^2(I;E))} + \|\delta(\sigma)\|^2_{L^2(\Omega;E)} \\
&\lesssim_E \|\xi_0\|^2_{L^2(\Omega;E)} + \|b\|^2_{L^2(\Omega;L^2(I;E))} + \|\sigma\|^2_{\mathbb{H}^{1,2}(L^2(I;\gamma(H,E)))},
\end{aligned}$$

and $\xi_t \in L^2(\Omega; E)$ is well defined, for any $t \in I$. □

As in [149], for any $t \in I$, we understand $\mathbb{1}_{[0,t]} : \mathscr{H} \to \mathscr{H}$ as a bounded and linear operator defined as

$$h \in \mathscr{H} \mapsto (\mathbb{1}_{[0,t]}h)(\cdot) \triangleq \mathbb{1}_{[0,t]}(\cdot)h(\cdot).$$

***Remark* 5.1.** *According to Lemma 5.1, we may regard $\mathbb{1}_{[0,t]}$ as a well defined operator on $\gamma(\mathscr{H}, \mathbb{H}^{2,2}(E))$, by setting $(\mathbb{1}_B\vartheta)h \triangleq \vartheta(\mathbb{1}_B h)$, for any $\vartheta \in \gamma(\mathscr{H}, \mathbb{H}^{1,2}(E))$.*

Proof of Lemma 5.7. Note that by linearity it is enough to prove that $\delta(\mathbb{1}_{[0,t]}\sigma) \in \mathbb{H}^{1,2}(E)$, for any $t \in I$, since $\xi_0 \in \mathbb{H}^{1,2}(E)$ and $b \in \mathbb{H}^{1,2}(L^2(I, E))$.

First, we may regard σ as an element of $\gamma(\mathscr{H}, \mathbb{H}^{2,2}(E))$, since from Lemma 5.4 we have $\sigma \in \mathbb{H}^{2,2}(\gamma(\mathscr{H}; E))$ and the space $\mathbb{H}^{2,2}(\gamma(\mathscr{H}, E))$ is isometric to $\gamma(\mathscr{H}, \mathbb{H}^{2,2}(E))$, (see [149], Theorem 2.9). Thus, according to Remark 5.1 we have that $\mathbb{1}_{[0,t]}\sigma \in \mathbb{H}^{2,2}(\gamma(\mathscr{H}, E)))$, for any $t \in I$, and hence that $\delta(\mathbb{1}_{[0,t]}\sigma) \in \mathbb{H}^{1,2}(E)$, by Proposition 5.2. □

Itô Formula. The following lemma introduces the notation for the trace operator $\mathrm{tr}(\cdot;\cdot)$.

Lemma 5.9. *For any for any $T \in \mathcal{L}(E, E^*)$ and $\vartheta \in \gamma(H, E)$, the series*

$$tr(T;\vartheta) \triangleq \sum_{n\geq 1} \langle T(\vartheta h_n), \vartheta h_n\rangle_E, \tag{5.36}$$

converges and its sum does not depend on the choice of the orthonormal basis $h_1, h_2, \ldots$ of H.

Proof. See, e.g., Lemma 2.3 in [31]. □

Let D be some Banach space. A function $\psi : I \times E \to D$ is said to be of class $\mathscr{C}^{1,2}$ if it is differentiable in the first variable and twice continuously Fréchet differentiable in the second variable and the functions ψ, $\nabla_k\psi$, for $k = 1, 2$, and $\nabla_2^2\psi$ are continuous on $I \times E$.

Theorem 5.5 (Itô formula)**.** *Let E and D be UMD Banach spaces. Moreover, let $\psi : I \times E \to D$ be a function of class $\mathscr{C}^{1,2}$ and ξ be the process defined by (6.2). Then, for any $t \in I$, the process $\nabla_2\psi(s, \xi_s)\sigma_s$, for $0 \leq s \leq t$, is stochastically integrable and the following identity holds a.s.*

$$\psi(t, \xi_t) = \psi(0, \xi_0) + \int_0^t a_s(\psi)ds + \int_0^t \nabla_2\psi(s, \xi_s)\sigma_s dW_s,$$

where

$$a_s(\psi) = \nabla_1\psi(s, \xi_s) + \nabla_2\psi(s, \xi_s)b_s + \frac{1}{2}tr(\nabla_2^2\psi(s, \xi_s); \sigma_s). \tag{5.37}$$

Proof. See, e.g., Theorem 2.4 in [31]. □

The following result provides a useful extension to the Itô formula.

Corollary 5.2. *Let E and D be UMD Banach spaces. Consider an adapted and strongly measurable E^*-valued stochastic process $b^* = \{b_t^* : t \in I\}$, that belongs to $L^2(\Omega; L^2(I; E^*))$. Let $\sigma^* = \{\sigma_t^* : t \in I\}$ be some adapted and H-strongly measurable*

$\mathcal{L}(H, E^*)$-valued process that belongs to $L^2(\Omega; L^2(I; \gamma(H, E^*)))$ and that is stochastic integrable with respect to W. Thus, let ξ be the process given by the identity (5.35) and set

$$\xi_t^* \triangleq \xi_0^* + \int_0^t b_s^* ds + \int_0^t \sigma_s^* dW_s. \tag{5.38}$$

Then, for any $t \in I$, we have

$$\begin{aligned}\langle \xi_t^*, \xi_t \rangle_E = \langle \xi_0^*, \xi_0 \rangle_E + \int_0^t \langle \xi_s^*, b_s \rangle_E + \langle b_s^*, \xi_s \rangle_E ds \\ + \int_0^t \langle \xi_s^*, \sigma_s \rangle_E + \langle \sigma_s^*, \xi_s \rangle_E + \int_0^t \sum_{n \geq 1} \langle \sigma_s^* h_n, \sigma_s h_n \rangle_E ds.\end{aligned}$$

Proof. See, e.g., Corollary 2.6 in [31]. □

CHAPTER 6

Portfolio Representation

In portfolio management, one encounters plenty of different situations characterized by the question of the efficient substitution of a portfolio of market securities with a simpler one which owns similar risk. Besides, it is reasonable to assume this simpler portfolio to meet certain contingent restrictions. We address this problem with the term of *portfolio representation*. For instance, one encounters this question when defining a hedging strategy subject to policy and budget constraints. Another instance of such a problem is obtained when one seeks to reduce the scale, and hence the complexity, of a specific portfolio for analysis and management purposes, without misrepresenting its inherent risk structure.

In this chapter, this problem is addressed by defining a reasonable notion of optimality based on a specific comparison criterion. More precisely, we present an approach consisting in the minimization of a certain risk functional, which gauges the average discrepancy between the original portfolio and any possible candidate among all the simpler portfolios representing such an exposure. In particular, the risk is gauged in terms of the fluctuation of the underlying risk factors within a given time horizon. We model the risk factors dynamics as a diffusion process in a certain Banach space and then we show two different formulations of this functional within the theory of the stochastic integration in UMD spaces developed by Van Neerven, Veraar and Weis [173, 174, 177]. The first formulation is shown in terms of the Malliavin derivative operator. In particular, the criterion we obtain turns out to be similar to the minimization approach suggested in [110] to address the problem of the optimal hedging of bond portfolios, in which a refined notion of duration is introduced by using Malliavin calculus for Gaussian random fields in the Hilbert space framework. The second formulation is obtained under further conditions on the model and it mainly involves the diffusive component of the discount curve dynamics.

Finally, we discuss how these arguments may apply to the fixed income market, by considering an infinite-dimensional dynamics governing the stochastic evolution of the price curve. We study the case in which an interest rate securities portfolio is represented by considering a portfolio composed by a single zero coupon bond. As a direct result, a refined notion of duration is recovered when minimizing the risk functional above.

We fix E to be a Banach space with type 2 and we write I to denote the unit interval on the real line. Furthermore, we fix a separable Hilbert space H and an H-cylindrical Wiener process W. The augmentation of the filtration generated by W is denoted by $\mathscr{G}^W$ and we fix $\mathscr{H} \triangleq L^2(I;H)$.

This chapter is based on the original work [79].

6.1 RISK FUNCTIONAL AND OPTIMIZATION

Here and in the sequel, we fix a strongly $\mathscr{G}_0^W$-measurable random variable $\xi_0 \in L^2(\Omega;E)$. Further, we consider an adapted and strongly measurable E-valued stochastic process $b = \{b_t : t \in I\}$ that belongs to $L^2(\Omega;L^2(I;E))$, and we fix an adapted and H-strongly measurable $\mathcal{L}(H,E)$-valued process $\sigma = \{\sigma_t : t \in I\}$ that belongs to $L^2(\Omega; L^2(I; \gamma(H, E)))$. Moreover, all the previous processes are assumed to satisfy the following conditions,

$$\xi_0 \in \mathbb{H}^{1,2}(E), \quad b \in \mathbb{H}^{1,2}(L^2(I; E)), \quad \sigma \in \mathbb{H}^{2,2}(L^2(I; \gamma(H, E))). \tag{6.1}$$

Then, we define the E-valued process $\xi \triangleq \{\xi_t : t \in I\}$ by setting

$$\xi_t \triangleq \xi_0 + \int_0^t b_s ds + \int_0^t \sigma_s dW_s, \qquad \text{for any } t \in I. \tag{6.2}$$

Discrepancy and P-sets. Let D be some UMD Banach space. Recall that a function $\psi : I \times E \to D$ is said to be of class $\mathscr{C}^{1,2}$ if it is differentiable in the first variable and twice continuously Fréchet differentiable in the second variable and the functions ψ, $\nabla_k \psi$, for $k = 1, 2$, and $\nabla_2^2 \psi$ are continuous on $I \times E$. Moreover, we shall say that ψ is of class $\mathscr{C}_b^{1,2}$ when in addition the following condition is satisfied

$$\|\nabla_2 \psi\|_\infty \triangleq \sup_{(t,x) \in I \times E} \|\nabla_2 \psi(t, x)\|_{\mathcal{L}(E,D)} < \infty. \tag{6.3}$$

Here and in the sequel, we write $\nabla_k \psi$ to denote the derivative of ψ with respect to the kth component, for any $k = 1, 2$.

Definition 6.1. We say that a function $\psi : I \times E \to D$ of class $\mathscr{C}^{1,2}$ is a *BS-function* relative to ξ, if the following condition holds true a.s.

$$\nabla_1 \psi(t, \xi_t) + \frac{1}{2} \mathrm{tr}(\nabla_2^2 \psi(t, \xi_t); \sigma_t) = 0, \quad \text{for any } t \in I. \tag{6.4}$$

The result below characterizes the dynamics of the process $\psi(t, \xi_t)$, for $t \in I$, when ψ is a BS-function relative to ξ.

Lemma 6.1. *Let $\psi : I \times E \to D$ be a function of class $\mathscr{C}^{1,2}$. If ψ is assumed to be a BS-function relative to ξ, then*

$$\psi(t, \xi_t) = \psi(0, \xi_0) + \int_0^t \nabla_2 \psi(s, \xi_s) b_s ds + \int_0^t \nabla_2 \psi(s, \xi_s) \sigma_s dW_s, \quad \textit{a.s., for any } t \in I. \tag{6.5}$$

Proof. Fix $t \in I$. Since the function ψ is assumed to be of class $\mathscr{C}^{1,2}$, Theorem 5.5 gives that the process $s \mapsto \nabla_2\psi(s,\xi_s)\sigma_s$, for $s \leq t$, is stochastically integrable and the following representation holds true,

$$\psi(t,\xi_t) = \psi(0,\xi_0) + \int_0^t a_s(\psi)ds + \int_0^t \nabla_2\psi(s,\xi_s)\sigma_s dW_s, \qquad \text{a.s.} \tag{6.6}$$

where, for any $s \leq t$, we set

$$a_s(\psi) \triangleq \nabla_1\psi(s,\xi_s) + \nabla_2\psi(s,\xi_s)b_s + \frac{1}{2}\mathrm{tr}(\nabla_2^2\psi(s,\xi_s);\sigma_s). \tag{6.7}$$

Thus, the result follows directly from (6.6) jointly with (6.7), since ψ is assumed to be a BS-function relative to ξ. $\square$

We now characterize certain classes of E^*-valued processes that we shall use later on.

Definition 6.2. By a *P-set* relative to ξ we understand any set Φ of E^*-valued and adapted processes $\phi \triangleq \{\phi_t : t \in I\}$ such that the following conditions are satisfies:

(i) For any $\phi \in \Phi$, one has that $\|\phi\|_\infty < \infty$, where we set

$$\|\phi\|_\infty \triangleq \inf\{C \geq 0 : \|\phi_t\|_{E^*} \leq C \text{ a.s. for any } t \in I\}.$$

(ii) For any $\phi \in \Phi$, there exists a function $\varphi : I \times E \to E^*$ of class $\mathscr{C}_b^{1,2}$, such that

$$\phi_t = \varphi(t,\xi_t), \qquad \text{a.s., for any } t \in I.$$

(iii) For any $\phi \in \Phi$, the following identity holds true a.s.,

$$\langle\phi_t,\xi_t\rangle_E = \langle\phi_0,\xi_0\rangle_E + \int_0^t \langle\phi_s,b_s\rangle_E ds + \int_0^t \langle\phi_s,\sigma_s\rangle_E dW_s, \tag{6.8}$$

Notice that in Definition 6.2, given any $\phi \in \Phi$, the variables $\langle\phi_t,\sigma_t\rangle_E$, for $t \in I$, form an adapted H-valued process. In this particular case, $\langle\cdot,\cdot\rangle_E$ is thus regarded as the H-valued dual pairing between $\gamma(H,E)$ and E^*. Besides, from inequality (5.5), one has that

$$\|\langle\phi_t,\sigma_t\rangle_E\|_H \leq \|\phi_t\|_{E^*}\|\sigma_t\|_{\gamma(H,E)}, \qquad \text{a.s., for any } t \in I. \tag{6.9}$$

Definition 6.3. Fix a P-set Φ relative to ξ and consider a function $f : I \times E \to \mathbb{R}$. For any $\phi \in \Phi$, the process $F(\phi) \triangleq \{F_t(\phi) : t \in I\}$ defined by

$$F_t(\phi) \triangleq f(t, \xi_t) - \langle \phi_t, \xi_t \rangle_E, \qquad \text{for any } t \in I, \tag{6.10}$$

is said to be the *discrepancy process* between f and ϕ relative to ξ.

If not otherwise specified, where a function $f : I \times E \to \mathbb{R}$ and a P-set Φ relative to ξ are fixed, for any $\phi \in \Phi$, we always write $F(\phi)$ to denote the discrepancy process between f and ϕ relative to ξ.

Lemma 6.2. *Let Φ be a P-set relative to ξ and consider a function $f : I \times E \to \mathbb{R}$. If f is of class $\mathscr{C}_b^{1,2}$, then one has that $F_t(\phi) \in L^2(\Omega)$, for any $t \in I$ and $\phi \in \Phi$, with*

$$\sup_{t \in I} \|F_t(\phi)\|^2_{L^2(\Omega)} < \infty.$$

Proof. First, notice that since $f(\cdot, 0)$ is assumed to be continuous on I, we have that

$$\|f(\cdot, 0)\|_\infty \triangleq \sup_{t \in I} |f(t, 0)| < +\infty.$$

Besides, for any $t \in I$, the following inequalities hold true a.s.

$$\begin{aligned} |f(t, \xi_t)|^2 &\leq \|\nabla_2 f\|^2_\infty \|\xi_t\|^2_E + |f(t, 0)|^2 \\ &\leq \|\nabla_2 f\|^2_\infty \|\xi_t\|^2_E + \|f(\cdot, 0)\|^2_\infty. \end{aligned}$$

Fix $\phi \in \Phi$, and note that $|\langle \phi_t, \xi_t \rangle_E| \leq \|\phi_t\|_{E^*} \|\xi_t\|_E$ a.s. for any $t \in I$. Thus, we obtain

$$\sup_{t \in I} \|F_t(\phi)\|^2_{L^2(\Omega)} \leq (\|\nabla_2 f\|^2_\infty + \|\phi\|^2_\infty) \sup_{t \in I} \|\xi_t\|^2_{L^2(\Omega; E)} + \|f(\cdot, 0)\|^2_\infty,$$

and the result holds true since the function f is assumed to be of class $\mathscr{C}_b^{1,2}$, jointly with the condition (i) in Definition 6.2 and Lemma 5.8. $\square$

Recall that for any random variable $X \in L^2(\Omega; E)$ that is differentiable in the Malliavin sense, we write DF to denote its Malliavin derivative. As discussed in §5.2, recall that the Malliavin derivative is regarded as a closed operator $D : L^2(\Omega; E) \to L^2(\Omega; \gamma(\mathscr{H}, E))$ and its domain in $L^2(\Omega; E)$ is denoted by $\mathbb{H}^{1,2}(E)$. Moreover, in the particular case when $E = \mathbb{R}$ we write $\mathbb{H}^{1,2} \triangleq \mathbb{H}^{1,2}(\mathbb{R})$. The following result will play a relevant role later on in this section.

Lemma 6.3. *Let Φ be a P-set relative to ξ and $f : I \times E \to \mathbb{R}$ a function of class $\mathscr{C}_b^{1,2}$. Then $F_t(\phi) \in \mathbb{H}^{1,2}$, for any $t \in I$ and $\phi \in \Phi$, with*

$$DF_t(\phi) = (\nabla_2 f(t, \xi_t) - \phi_t) D\xi_t, \qquad a.s. \tag{6.11}$$

It is worth to be highlighted that, for any $t \in I$, identity (6.11) is understood as

$$DF_t(\phi) = \nabla_2 f(t, \xi_t) D\xi_t - \langle \phi_t, D\xi_t \rangle_E, \qquad \text{a.s.} \tag{6.12}$$

In particular, since $D\xi_t \in L^2(\Omega; \gamma(\mathscr{H}, E))$, for any $t \in I$, in identity (6.12) we understand $\langle \cdot, \cdot \rangle_E$ as the $\mathscr{H}$-valued dual pairing between $\gamma(\mathscr{H}, E)$ and E^*.

Proof of Lemma 6.3. Here and throughout, we fix $t \in I$. First of all, since the function f is assumed to be of class $\mathscr{C}_b^{1,2}$ and $\xi_t \in \mathbb{H}^{1,2}(E)$ thanks to Lemma 5.7, we have that Proposition 5.9 applies and and we get $f(t, \xi_t) \in \mathbb{H}^{1,2}$, with

$$Df(t, \xi_t) = \nabla_2 f(t, \xi_t) D\xi_t, \qquad \text{a.s.} \tag{6.13}$$

Fix now $\phi \in \Phi$. According to the assumption (ii) in Definition 6.2, there exists a function $\varphi : I \times E \to E^*$ of class $\mathscr{C}_b^{1,2}$ such that the identity $\phi_t = \varphi(t, \xi_t)$ holds true a.s. Thus, according to Theorem 5.5, we obtain that the process $\nabla_2 \varphi(s, \xi_s) \sigma_s$, for $s \leq t$, is stochastically integrable and the following representation holds a.s.,

$$\varphi(t, \xi_t) = \varphi(0, \xi_0) + \int_0^t a_s(\varphi) ds + \int_0^t \chi_s(\varphi) dW_s,$$

where, for $s \leq t$, we set,

$$\begin{aligned} a_s(\varphi) &\triangleq \nabla_1 \varphi(s, \xi_s) + \nabla_2 \varphi(s, \xi_s) b_s + \frac{1}{2} \mathrm{tr}(\nabla_2^2 \varphi(s, \xi_s); \sigma_s); \\ \chi_s(\varphi) &\triangleq \nabla_2 \varphi(s, \xi_s) \sigma_s. \end{aligned}$$

Hence, given any orthonormal basis $h_1, h_2, \dots$ in H, the extension of the Itô's formula in Corollary 5.2 gives

$$\begin{aligned} \langle \varphi(t, \xi_t), \xi_t \rangle_E = \langle \varphi(0, \xi_0), \xi_0 \rangle_E &+ \int_0^t \langle \varphi(s, \xi_s), b_s \rangle_E ds + \int_0^t \langle \varphi(s, \xi_s), \sigma_s \rangle_E dW_s \\ &+ \int_0^t \langle a_s, \xi_s \rangle_E ds + \int_0^t \langle \chi_s, \xi_s \rangle_E dW_s + \int_0^t \sum_{n \geq 1} \langle \chi_s h_n, \sigma_s h_n \rangle_E ds, \qquad \text{a.s.} \end{aligned} \tag{6.14}$$

As a direct consequence, the condition (6.8) jointly with the identity (6.14) implies that $\chi_t = 0$ a.s. In particular

$$\nabla_2 \varphi(t, \xi_t) = 0, \qquad \text{a.s.} \tag{6.15}$$

On the other hand, Proposition 5.9 applies and it implies that $\varphi(t, \xi_t) \in \mathbb{H}^{1,2}(E^*)$, with

$$D\varphi(t, \xi_t) = \nabla_2 \varphi(t, \xi_t) D\xi_t, \qquad \text{a.s.} \tag{6.16}$$

Thus, jointly with the identity (6.15), we get

$$D\varphi(t, \xi_t) = 0, \qquad \text{a.s.} \tag{6.17}$$

On the other hand, the product rule for the Malliavin derivative stated in Proposition 5.9 jointly with (6.17) implies that

$$\begin{aligned} D\langle \varphi(t, \phi_t), \xi_t \rangle_E &= \langle D\varphi(t, \xi_t), \xi_t \rangle_E + \langle \varphi(t, \xi_t), D\xi_t \rangle_E, \quad \text{a.s.} \\ &= \langle \varphi(t, \xi_t), D\xi_t \rangle_E \quad \text{a.s.} \end{aligned} \tag{6.18}$$

Then, from the inequality (5.5) and by means of the natural identification $\gamma(\mathscr{H}, \mathbb{R}) = \mathscr{H}$, we obtain that

$$\begin{aligned} \|D\langle \varphi(t, \xi_t), \xi_t \rangle_E\|^2_{L^2(\Omega;\mathscr{H})} &= \mathbb{E}\{\|\langle \varphi(t, \xi_t), D\xi_t \rangle_E\|^2_{\mathscr{H}}\} \\ &\leq \mathbb{E}\{\|\varphi(t, \xi_t)\|^2_{E^*} \|D\xi_t\|^2_{\gamma(\mathscr{H},E)}\} \\ &\leq \|\phi\|^2_\infty \|D\xi_t\|^2_{L^2(\Omega,\gamma(\mathscr{H},E))} \end{aligned}$$

Thus, since $\xi_t \in \mathbb{H}^{1,2}(E)$ due to Lemma 5.7, we have $\langle \phi_t, \xi_t \rangle_E \in \mathbb{H}^{1,2}$ and hence $F_t(\phi) \in \mathbb{H}^{1,2}$. Finally, the identity (6.11) follows by the linearity of the Malliavin derivative from equations (6.13) and (6.18). $\square$

Risk functional. Later on, we introduce a particular functional induced by the discrepancy process (6.10). In the theory of portfolio representation, this process admits a natural interpretation that we shall discuss in the sequel.

Definition 6.4. Let Φ be a P-set relative to ξ. Given a function $f : I \times E \to \mathbb{R}$ of class $\mathscr{C}_b^{1,2}$, we refer to the functional $\mathcal{F} : \Phi \to \mathbb{R}$ defined by setting

$$\mathcal{F}(\phi) \triangleq \int_I \mathbb{E}\{|F_t(\phi) - \mathbb{E}F_t(\phi)|^2\}dt, \qquad \text{for any } \phi \in \Phi, \tag{6.19}$$

as the *risk functional* relative to ξ induced by f over Φ.

If not otherwise specified, where a function $f : I \times E \to \mathbb{R}$ and a P-set Φ relative to ξ are fixed, we always use $\mathcal{F}$ to denote the risk functional relative to ξ induced by f over Φ. The following theorem implies that the risk functional (6.19) admits two different equivalent representations.

Theorem 6.1. *Let Φ be a P-set relative to ξ and $f : I \times E \to \mathbb{R}$ a function of class $\mathscr{C}_b^{1,2}$. Then,*

(i) The functional $\mathcal{F}$ admits the following representation

$$\mathcal{F}(\phi) = \int_I \mathbb{E}\left\{ \int_0^t \left\| \mathbb{E}\{(\nabla_2 f(t, \xi_t) - \phi_t) D_s \xi_t | \mathscr{G}_s^W\} \right\|_H^2 ds \right\} dt, \qquad \textit{for any } \phi \in \Phi. \tag{6.20}$$

(ii) In the special case when f is assumed to be a BS-function relative to ξ and $b_t = 0$ a.s., for any $t \in I$, the functional $\mathcal{F}$ boils down to

$$\mathcal{F}(\phi) = \mathbb{E}\left\{ \int_I \left\| (\nabla_2 f(t, \xi_t) - \phi_t)\sigma_t \right\|_H^2 (1-t) dt \right\}, \qquad \textit{for any } \phi \in \Phi. \tag{6.21}$$

It is worth to be noted that, since the process σ takes values in $\gamma(H, E)$, the right hand side of the identity (6.21) is understood as follows

$$(\nabla_2 f(s, \xi_s) - \phi_s)\sigma_s \triangleq \nabla_2 f(s, \xi_s)\sigma_s - \langle \phi_s, \sigma_s \rangle_E,$$

where $\langle \cdot, \cdot \rangle_E$ denotes the H-valued dual pairing between $\gamma(H, E)$ and E^*. The proof of Theorem 6.1 is based on the following lemma.

Lemma 6.4. *Let Φ be a P-set relative to ξ and $f : I \times E \to \mathbb{R}$ a function of class $\mathscr{C}_b^{1,2}$. In the particular case when f is assumed to be a BS-function relative to ξ and $b_t = 0$ a.s., for any $t \in I$, we have a.s.*

$$F_t(\phi) = F_0(\phi) + \int_0^t (\nabla_2 f(s, \xi_s) - \phi_s)\sigma_s dW_s, \tag{6.22}$$

for any $t \in I$ and $\phi \in \Phi$.

Proof of Lemma 6.4. Fix $t \in I$. Since the function f is assumed to be a BL-function relative to ξ, Lemma 6.1 gives that

$$f(t, \xi_t) = f(0, \xi_0) + \int_0^t \nabla_2 f(s, \xi_s) b_s ds + \int_0^t \nabla_2 f(s, \xi_s)\sigma_s dW_s, \qquad \text{a.s.}$$

Thus, for any $\phi \in \Phi$, from condition (6.8) we get that the variable $F_t(\phi)$, for $t \in I$, admits the following representation,

$$F_t(\phi) = F_0(\phi) + \int_0^t (\nabla_2 f(s, \xi_s) - \phi_s) b_s ds + \int_0^t (\nabla_2 f(s, \xi_s) - \phi_s)\sigma_s dW_s, \quad \text{a.s.} \tag{6.23}$$

Besides, since $b_t = 0$ a.s., the representation (6.23) boils down to the identity (6.22). $\square$

Proof of Theorem 6.1. First, we prove the statement (i). Since f is assumed to be of class $\mathscr{C}_b^{1,2}$, from Lemma 6.3 we get that $F_t(\phi) \in \mathbb{H}^{1,2}$, for any $\phi \in \Phi$ and $t \in I$, and that

$$DF_t(\phi) = (\nabla_2 f(t, \xi_t) - \phi_t) D\xi_t, \qquad \text{a.s.} \tag{6.24}$$

Thus, fix $t \in I$ and notice that, since the variable $F_t(\phi)$ is $\mathscr{G}_t^W$-measurable, Clarke-Ocone formula in Theorem 5.4 implies that,

$$F_t(\phi) - \mathbb{E}F_t(\phi) = \int_0^t \mathbb{E}\{D_s F_t(\phi) | \mathscr{G}_s^W\} dW_s \qquad \text{a.s.} \tag{6.25}$$

On the other hand, a direct application of Lemma 5.5 leads to

$$\mathbb{E}\left\{\left| \int_0^t \mathbb{E}\{D_s F_t(\phi) | \mathscr{G}_s^W\} dW_s \right|^2\right\} = \mathbb{E}\left\{ \int_0^t \|\mathbb{E}\{D_s F_t(\phi) | \mathscr{G}_s^W\}\|_H^2 ds \right\}. \tag{6.26}$$

Then, when recasting the identity (6.26) in terms of the representation (6.24), jointly with the identity (6.25), we obtain

$$\mathbb{E}\{|F_t(\phi) - \mathbb{E}F_t(\phi)|^2\} = \mathbb{E}\bigg\{\int_0^t \|\mathbb{E}\{(\nabla_2 f(t,\xi_t) - \phi_t)D_s\xi_t|\mathscr{G}_s^W\}\|_H^2 ds\bigg\}. \tag{6.27}$$

As a result, when integrating both the sides of the identity (6.27) with respect to the variable $t \in I$, we obtain the representation (6.20).

We now prove the statement (ii). For this purpose, fix $\phi \in \Phi$ and notice that $\mathbb{E}F_t(\phi) = \mathbb{E}F_0(\phi)$ a.s., for any $t \in I$. Next, since f is assumed to be a BL-function relative to ξ and $b_t = 0$, a.s. for any $t \in I$, Lemma 6.4 gives that the following representation holds true a.s.

$$F_t(\phi) - \mathbb{E}F_t(\phi) = F_0(\phi) - \mathbb{E}F_0(\phi) + \int_0^t (\nabla_2 f(s,\xi_s) - \phi_s)\sigma_s dW_s, \qquad \text{for any } t \in I. \tag{6.28}$$

Then, let $p(x) = |x|^2$, for any $x \in \mathbb{R}$. According to the representation (6.28), for any $t \in I$, a direct application of Theorem 5.5 leads to

$$\begin{aligned}&|F_t(\phi) - \mathbb{E}F_t(\phi)|^2 = |F_0(\phi) - \mathbb{E}F_0(\phi)|^2 \\ &+ \frac{1}{2}\int_0^t \operatorname{tr}(\nabla^2 p(F_s(\phi) - \mathbb{E}F_s(\phi)); (\nabla_2 f(s,\xi_s) - \phi_s)\sigma_s) ds + \int_0^t \kappa_s(\phi) dW_s, \ \text{ a.s.,}\end{aligned} \tag{6.29}$$

where, for any $s \leq t$, we set

$$\kappa_s(\phi) \triangleq 2(F_s(\phi) - \mathbb{E}F_s(\phi))(\nabla_2 f(s,\xi_s) - \phi_s)\sigma_s.$$

Let $h_1, h_2, ...$ be an orthonormal basis of H and note that $\nabla^2 p(x) = 2$, for any $x \in \mathbb{R}$. Then, by definition of the trace operator $\operatorname{tr}(\cdot;\cdot)$, we obtain,

$$\begin{aligned}\operatorname{tr}(\nabla_2^2 p(F_s(\phi) - \mathbb{E}F_s(\phi)); (\nabla_2 f(s,\xi_s) - \phi_s)\sigma_s) &= 2\sum_{n\geq 1} (\nabla_2 f(s,\xi_s) - \phi_s)\sigma_s h_n)^2 \\ &= 2\|(\nabla_2 f(s,\xi_s) - \phi_s)\sigma_s\|_H^2.\end{aligned}$$

Notice that $\mathbb{E}|F_0(\phi) - \mathbb{E}F_0(\phi)|^2 = 0$, since $F_0(\phi)$ is $\mathscr{G}_0^W$-measurable, and hence $F_0(\phi) = \mathbb{E}F_0(\phi)$ a.s. Thus, from the identity (6.29) we get

$$\mathbb{E}\{|F_t(\phi) - \mathbb{E}F_t(\phi)|^2\} = \mathbb{E}\bigg\{\int_0^t \|(\nabla_2 f(s,\xi_s) - \phi_s)\sigma_s\|_H^2 ds\bigg\}. \tag{6.30}$$

When integrating in both the sides of the identity (6.30) with respect to $t \in I$, we obtain the representation (6.21). $\square$

Next, we consider the following notion of optimality that we state in general terms.

Definition 6.5. Let Υ be some set and consider a functional $\mathcal{G} : \Upsilon \to \mathbb{R}$. We call $\mathcal{G}$-*optimal* any element $\upsilon^* \in \Upsilon$ that verifies the following inequality

$$\mathcal{G}(\upsilon^*) \leq \mathcal{G}(\upsilon), \qquad \text{for any } \upsilon \in \Upsilon. \tag{6.31}$$

Discrete representation. Fix a P-set Φ relative to ξ and a function $f : I \times E \to \mathbb{R}$ of class $\mathscr{C}_b^{1,2}$. In the particular case when f is a BS-function relative to ξ, the statement (ii) in Theorem 6.1 implies that that a process $\phi^* \in \Phi$ is $\mathcal{F}$-optimal if it minimizes the functional

$$\mathcal{F}(\phi) = \mathbb{E}\left\{ \int_0^1 \|(\nabla_2 f(t, \xi_t) - \phi_t)\sigma_t\|_H^2 (1-t)dt \right\}, \qquad \text{for any } \phi \in \Phi. \tag{6.32}$$

Besides, in the particular case when fixing $x_1^*, ..., n_n^* \in E^*$ and assuming that for any $\phi \in \Phi$ there exists $\beta = (\beta_1, ..., \beta_n) \in \mathbb{R}^n$ such that the following representation holds true a.s.

$$\phi_t = \sum_{i=1}^n \beta_i x_i^*, \qquad \text{for any } t \in I, \tag{6.33}$$

the minimization of the functional (6.32) turns out to be analytically tractable, since it boils down to a quadratic optimization problem. Indeed, when identifying any process $\phi \in \Phi$ with the element $\beta \in \mathbb{R}^n$ that satisfies the identity (6.33), *mutatis mutandis* the functional (6.32) may be recast as follows

$$\mathcal{F}(\beta) = \sum_{ij=1}^n A_{ij}\beta_i\beta_j - 2\sum_{i=1}^n B_i\beta_i + C, \qquad \text{for any } \beta \in \mathbb{R}^n,$$

where, for any $i, j = 1, ..., n$ we set

$$\begin{aligned}
A_{ij} &= \mathbb{E}\left\{ \int_0^1 \left\langle \langle x_i^*, \sigma_t \rangle_E, \langle x_j^*, \sigma_t \rangle_E \right\rangle_H (1-t)dt \right\}, \\
B_i &= \mathbb{E}\left\{ \int_0^1 \left\langle \nabla_2 f(t, \xi_t)\sigma_t, \langle x_i^*, \sigma_t \rangle_E \right\rangle_H (1-t)dt \right\}, \\
C &= \mathbb{E}\left\{ \int_0^1 \|\nabla_2 f(t, \xi_t)\sigma_t\|_H^2 (1-t)dt \right\}.
\end{aligned}$$

As a result, in the special case when the symmetric matrix $A = (A_{ij})_{ij}$ turns out to be positive definite, there is a unique $\mathcal{F}$-optimal element $\beta^* \in \Phi$. Furthermore, when defining $B \triangleq (B_1, ..., B_n) \in \mathbb{R}^n$, the element $\beta^* \in \Phi$ is obtained as the solution the following n-dimensional inverse problem,

$$A\beta^* = B.$$

6.2 THE PORTFOLIO REPRESENTATION PROBLEM

In this section, we discuss how the results presented in Section 6.1 may be used to address the problem of substituting a financial exposure by some constrained portfolio, without misrepresenting its performance in terms of the related inherent risk structure. Here and in the sequel, we always assume that the risk factors of a given portfolio are represented by tradable observables. These include the market price of a stock or a commodity itself or some major market benchmark assessing the value of an entire class of securities, such as the interest rate term structure when dealing with the fixed income market. In this respect, we only refer to the market risk that affects the financial exposure. No other types of risk are considered.

Risk factors and financial exposure. Any element of E represents the overall discounted value of the risk factors at a certain time and we regard I as the reference time interval. For sake of simplicity, we suppose I to define the period of one year, and any of its fractions to be assessed according to a certain day count convention. To fix the ideas, when E is chosen to be the Euclidean space of some finite dimension

$n \in \mathbb{N}$, then the components of any element $x = (x_1, ..., x_n) \in E$ may represent the discounted market prices of n assets at a certain time. Besides, one may chose E to be some space of continuous curves and interpret any of its elements as the structure of the discounted price curve at a certain time.

We suppose the process (6.2) to provide a dynamics for the overall discounted value of the risk factors. Moreover, we regard any function $f : I \times E \to \mathbb{R}$ of class $\mathscr{C}_b^{1,2}$ as a fixed financial exposure and we understand the variable $f(t, \xi_t)$ as its discounted value at any time $t \in I$. For a fixed P-set Φ relative to ξ, we will interpret any $\phi \in \Phi$ as the dynamics of a certain portfolio of risk factors managed by the trader. Hence, we understand the variable $\langle \phi_t, \xi_t \rangle_E$ as its discounted value at time $t \in I$. Notice that this term depends on both the overall discounted value of the risk factors ξ_t and the portfolio composition ϕ_t chosen by the investor. The assumption (ii) in Definition 6.2 is not merely technical, since the strategy considered by a rational investor at a certain time should depend on the evolution of the reference market. On the other hand, for our applications, the identity (iii) in Definition 6.2 generalizes the well-known self-financing condition.

We regard the risk functional $\mathcal{F}$ relative to ξ induced by f over Φ as the error that occurs when substituting the exposure represented by f with some portfolio within Φ. Such an error is assessed in terms of the average changes of the difference between the exposure and the selected portfolio, due to the fluctuation of the underling risk factors within the period I. In this respect, a portfolio $\phi^* \in \Phi$ turns out to be $\mathcal{F}$-optimal when it provides the best representation of the inherent risk of the exposure f, among all the possible choices within Φ.

Theorem 6.1 shows two different formulations of the functional $\mathcal{F}$ in terms of the operator $\nabla_2 f$, which gauges the sensitivity of the financial exposure to little variations of the underlying risk factors. In this respect, the optimization of the functional $\mathcal{F}$ may be understood as a notion of portfolio immunization via sensitivities analysis.

Market prices. We say that a integrable process $X \triangleq \{X_t : t \in I\}$ is a *market price* if

$$\mathbb{E}X_t = \mathbb{E}X_0, \qquad \text{for any } t \in I.$$

Notice that the process ξ given by the identity (6.2) is a market price when assuming that $b_t = 0$ a.s., for any $t \in I$.

On the other hand, it is worth to be noted that in the particular case when f is assumed to be a BS-function relative to ξ and $b_t = 0$ a.s., for any $t \in I$, Lemma 6.1 assures that

$$\mathbb{E}f(t, \xi_t) = \mathbb{E}f(0, \xi_0), \qquad \text{for any } t \in I.$$

Thus, when ξ is a market price and the function f turns out to be a BS-function relative to ξ, one has that also the process $f(t, \xi_t)$, for any $t \in I$, is a market price. As a consequence, we may regard Definition 6.1 as a risk-free condition.

On the other hand, under the same hypothesis Lemma 6.4 leads to

$$\mathbb{E}F_t(\phi) = \mathbb{E}F_0(\phi), \qquad \text{for any } t \in I \text{ and } \phi \in \Phi.$$

Hence, in the particular case when $f(0, \xi_0) = \langle \phi_0, \xi_0 \rangle_E$ a.s., for any $\phi \in \Phi$, one has $F_t(\phi) = 0$ a.s. for any $t \in I$ and $\phi \in \Phi$, and the risk functional $\mathcal{F}$ induced by f over Φ boils down to

$$\mathcal{F}(\phi) = \|F(\phi)\|^2_{L^2(\Omega \times I)}, \qquad \text{for any } \phi \in \Phi.$$

This means that when the financial exposure is market valued and it is perfectly hedged by any portfolio $\phi \in \Phi$ at time $t = 0$, the $\mathcal{F}$-optimal portfolio $\phi^* \in \Phi$ is the one that minimizes the average squared discrepancy with the exposure over time, based on the underline risk factors fluctuation.

6.3 CONSTRAINED HEDGING AND RESIDUAL RISK

Let $f : I \times E \to \mathbb{R}$ be a function of class $\mathscr{C}_b^{1,2}$ and fix a P-set Φ relative to ξ. As a consequence of Theorem 6.1, if there exists a process $\phi^* \in \Phi$ such that the following identity holds true a.s.

$$\phi_t^* = \nabla_2 f(t, \xi_t), \qquad \text{for any } t \in I, \tag{6.34}$$

then ϕ^* turns out to be $\mathcal{F}$-optimal, since $\mathcal{F}(\phi^*) = 0$ in this case. This fact is consistent with the sensitivity-based hedging approach for portfolio immunization. We discuss this interpretation later on.

Sensitivity-based hedging. Assume E to coincide with the Euclidean space $\mathbb{R}^2$. Hence, write $\langle \cdot, \cdot \rangle_E$ to denote the standard Euclidean product on it and let $e = (e_1, e_2)$ be its canonical basis. On the other hand, we assume H to coincide with $\mathbb{R}$ and thus we regard the H-cylindrical process W as a standard one-dimensional Brownian motion.

Given this framework, note that the space $\gamma(H, E)$ boils down to E itself. Thus, we may represent the process $\sigma \in L^2(\Omega, L^2(I, E))$ in terms of its components by writing $\sigma = (\sigma_1, \sigma_2)$, where, for any $i = 1, 2$, the real valued process $\sigma_i = \{\sigma_{i,t} : t \in I\}$ is defined as follows,

$$\sigma_{i,t} \triangleq \langle e_i, \sigma_t \rangle_E, \qquad \text{for any } t \in I.$$

Similarly, we write $\xi_i \triangleq \{\xi_{i,t} : t \in I\}$, for $i = 1, 2$, to denote the components of the process ξ that are defined by setting

$$\xi_{i,t} \triangleq \langle e_i, \xi_t \rangle_E, \qquad \text{for } t \in I. \tag{6.35}$$

Throughout, we assume that a.s. $b_t = 0$ and $\sigma_{2,t} = 0$, for any $t \in I$. Moreover, for sake of simplicity, we set $\xi_{2,0} = 1$ a.s. Then, since for $i = 1, 2$ the following equality a.s. holds true due to Theorem 5.2,

$$\left\langle e_i, \int_0^t \sigma_s dW_s \right\rangle_E = \int_0^t \langle e_i, \sigma_s \rangle_E dW_s, \qquad \text{for any } t \in I,$$

we obtain the following representation for the dynamics of the components given in (6.35),

$$\xi_{1,t} = \xi_{1,0} + \int_0^t \sigma_{1,s} dW_s, \quad \text{a.s.}, \qquad \xi_{2,t} = 1, \quad \text{a.s.} \tag{6.36}$$

We fix a function $f : I \times E \to \mathbb{R}$ of class $\mathscr{C}_b^{1,2}$ and a P-set Φ relative to ξ. For any process $\phi \in \Phi$, we regard the variable ϕ_t, for $t \in I$, in terms of its components defined for $i = 1, 2$ by the following identity,

$$\phi_{i,t} \triangleq \langle e_i, \phi_t \rangle_E, \qquad \text{for any } t \in I.$$

Proposition 6.1. *Let Φ be a P-set relative to ξ. Assume f to be a BL-function relative to ξ. If the components of the process ξ are given by (6.36), then*

$$\mathcal{F}(\phi) = \mathbb{E}\left\{ \int_I |(\partial_{x_1} f(t, \xi_{1,t}, \xi_{2,t}) - \phi_{1,t})\sigma_{1,t}|^2 (1-t)dt \right\}, \qquad \textit{for any } \phi \in \Phi. \quad (6.37)$$

Proof. Notice that for any $t \in I$ and $x = (x_1, x_2) \in E$, we may regard $\nabla_2 f(t, x)$ as an element of E with components,

$$\nabla_2 f(t, x) = (\partial_{x_1} f(t, x_1, x_2), \partial_{x_2} f(t, x_1, x_2)),$$

and thus we may set

$$\nabla_2 f(t, x) y \triangleq \langle \nabla_2 f(t, x), y \rangle_E, \qquad \text{for any } y \in E.$$

Since according to the identities (6.36) we have $b_t = 0$ a.s., for any $t \in I$, and f is a BL-function relative to ξ, then the result follows directly from the statement (ii) in Theorem 6.1, by noting that $\sigma_{2,t} = 0$ a.s., for any $t \in I$. □

We may understand the first component of the process ξ as a risk-neutral dynamics for the discounted price of some risky asset. Besides, we regard its second component as a risk-neutral model for the discounted value of the bank account. The function f represents an European contingent claim written on the risky asset and the P-set Φ relative to ξ stands for the entire class of the hedging portfolios. It is worth to be noted that in the particular case when there exists a process $\phi^* \in \Phi$ such that the following equality holds true a.s.,

$$\phi^*_{1,t} = \partial_{x_1} f(t, \xi_t), \qquad \text{for any } t \in I, \quad (6.38)$$

then the process ϕ^* turns out to be $\mathcal{F}$-optimal, since $\mathcal{F}(\phi^*) = 0$ thanks to the representation (6.37). In particular, the identity (6.38) corresponds to the so-called delta-hedging condition for the contingent claim f.

Assume that the first component of the process $\phi^* \in \Phi$ satisfies the identity (6.38). Note that, if the following condition holds a.s.

$$f(0, \xi_0) = \langle \phi^*_0, \xi_0 \rangle_E, \quad (6.39)$$

then $F_0(\phi^*) = 0$ a.s. The identity (6.39) implies that the financial exposure is perfectly hedged by the portfolio $\phi^* \in \Phi$ at time $t = 0$.
In this particular case, Lemma 6.4 gives that $F_t(\phi^*) = 0$ a.s., for any $t \in I$, and the second component of the process ϕ^* is thus implicitly determined by the following identity

$$\phi^*_{2,t} = f(t, \xi_t) - \phi^*_{1,t}\xi_{1,t} \qquad \text{a.s., for any } t \in I.$$

Moreover, it is worth to be noted that within the present setup the identity (6.5) boils down to

$$\nabla_1 f(t, \xi_{1,t}, \xi_{2,t}) + \frac{1}{2}\partial_{x_1x_1} f(t, \xi_{1,t}, \xi_{2,t})\sigma^2_{1,t} = 0, \quad \text{a.s., for any } t \in I,$$

which corresponds to a Black-Scholes type equation, whereby the risk-free rate is set to be null at any time.

Correlation and residual risk. Throughout, we analyse the case when the set Φ is defined in such a way that the condition (6.34) may not be fulfilled. This case turns out to be appealing when the financial exposure and the replication portfolio actually depends upon two different but correlated risk factors. To see this, assume both E and H to coincide with the Euclidean space $\mathbb{R}^2$. Hence, write $\langle\cdot,\cdot\rangle_E$ to denote the standard Euclidean product on it and let $e = (e_1, e_2)$ be its canonical basis. Further, fix a constant $\varrho \in (-1, 1)$ and assume the inner product $\langle\cdot,\cdot\rangle_H$ on H to be defined in such a way that $\langle e_i, e_j\rangle_H = 1$, if $i = j$, and $\langle e_i, e_j\rangle_H = \varrho$ otherwise. Let $f : I \times E \to \mathbb{R}$ be the function defined by

$$f(t, x) = \langle e_1, x\rangle_E, \qquad \text{for any } t \in I \text{ and } x \in E. \tag{6.40}$$

Notice that f is of class $\mathscr{C}^{1,2}_b$, with $\nabla_2 f(t, x) = e_1$, for any $t \in I$ and $x \in E$. Let Φ be a P-set relative to ξ and assume that for any $\phi \in \Phi$ there exists a real-valued process $\phi_2 \triangleq \{\phi_{2,t} : t \in I\}$ such that the following identity holds true

$$\phi_t = \phi_{2,t}e_2, \qquad \text{a.s., for any } t \in I. \tag{6.41}$$

Hence, we identify any $\phi \in \Phi$ with the process ϕ_2 that satisfies the identity (6.41).

Proposition 6.2. *Let f be the function given by the identity (6.40) and Φ a P-set relative to ξ that verifies the condition (6.41). If $b_t = 0$ a.s., for any $t \in I$, then*

$$\mathcal{F}(\phi) = \mathbb{E}\bigg\{ \int_I \|\langle e_1, \sigma_t\rangle_E - \phi_{2,t}\langle e_2, \sigma_t\rangle_E\|_H^2 (1-t)dt \bigg\}, \qquad \textit{for any } \phi \in \Phi. \tag{6.42}$$

Proof. First of all, notice that $\nabla_1 f(t,x) = 0$ and $\nabla_2^2 f(t,x) = 0$, for any $t \in I$ and $x \in I$. Hence, the function f appears to be a BS-function relative to ξ. Moreover, note that

$$\nabla_2 f(t,\xi_t)\sigma_t = \langle e_1, \sigma_t\rangle_E, \qquad \text{for any } t \in I.$$

and for any $\phi \in \Phi$, and that

$$\langle \phi_t, \sigma_t\rangle_E = \phi_{2,t}\langle e_2, \sigma_t\rangle_E, \qquad \text{for any } t \in I.$$

Then, the result follows directly from Theorem 6.1. □

Since it is not possible to find a version of the process $\nabla_2 f(t,\xi_t)$, for $t \in I$, that belongs to Φ, then the condition (6.34) may not be recovered. Moreover, the following result provides an explicit characterization of a $\mathcal{F}$-optimal process in this case.

Proposition 6.3. *Let f be the function given by the identity (6.40) and Φ a P-set relative to ξ that verifies the condition (6.41). If $b_t = 0$ a.s., for any $t \in I$, then the process $\phi^* \in \Phi$ defined by*

$$\phi^*_{2,t} = \varrho(\sigma_{1,t}/\sigma_{2,t}), \qquad \textit{for any } t \in I, \tag{6.43}$$

turns out to be $\mathcal{F}$-optimal, where for $i = 1,2$ the real valued process $\sigma_i \triangleq \{\sigma_{i,t} : t \in I\}$ is such that

$$\langle e_i, \sigma_t\rangle_E = \sigma_{i,t} e_i, \qquad \textit{for any } t \in I. \tag{6.44}$$

Proof. Notice that Proposition 6.2 applies and the functional $\mathcal{F}$ admits the representation given in (6.42). On the other hand, a direct computation shows for any $t \in I$, one has

$$\|\langle e_1, \sigma_t\rangle_E - \phi_{2,t}\langle e_2, \sigma_t\rangle_E\|_H^2 = \sigma_{1,t}^2 - 2\varrho\phi_{2,t}\sigma_{1,t}\sigma_{2,t} + \phi_{2,t}^2\sigma_{2,t}^2, \tag{6.45}$$

which consists in a quadratic form in terms of $\phi_{2,t}$, that attends its minimum at

$$\phi_{2,t} = \varrho(\sigma_{1,t}/\sigma_{2,t}).$$

Thus, when letting t run over I, the process $\phi^* \in \Phi$ defined by identity (6.43) satisfies $\mathcal{F}(\phi^*) \leq \mathcal{F}(\phi)$, for any $\phi \in \Phi$, and hence it turns out to be $\mathcal{F}$-optimal. $\square$

Notice that the H-Wiener process W may be understood as a 2-dimensional Wiener process, whose components W_i, for $i = 1, 2$, are defined by

$$W_{i,t} \triangleq W_t e_i, \qquad \text{for any } t \in I, \tag{6.46}$$

with ϱ as their instantaneous correlation, since

$$\mathbb{E}\{W_t e_1 \cdot W_t e_2\} = \langle e_1, e_2 \rangle_H t = \varrho t, \qquad \text{for any } t \in I.$$

Thus, we may understand the components of the process ξ as the dynamics of the discounted prices of two correlated risky assets. In this case, a perfect hedge may not be recovered. Indeed, notice that given $\phi^* \in \Phi$ as defined by the identity (6.43), the quantity $\mathcal{F}(\phi^*)$ is strictly positive for $\varrho < 1$ and it vanishes for $\varrho = 1$, which is the particular case when the two assets appear to be completely correlated. Then, we may interpret $\mathcal{F}(\phi^*)$ as the *residual hedging risk*.

6.4 INTEREST RATES SECURITIES PORTFOLIOS

In this section we take a close look at the problem of seeking an optimal representation of some fixed income portfolio by considering a portfolio of zero coupon bonds. Recall that a zero coupon bond is a contract that pays one unit of a certain currency at some maturity future date. We make the assumption to deal with idealized bonds that are unaffected by credit risk. i.e. the payment at the maturity date is made by the issuer of the bond.

We fix $\mathscr{T} \triangleq (1, +\infty)$. Thus, we suppose that there exists a market valued bond maturing at any future time $T \in \mathscr{T}$, and we write $p_t(T)$ to denote its risk-neutral discounted price at time $t \in I$. Moreover, we refer to the function $T \in \mathscr{T} \mapsto p_t(T)$ as the discounted price curve at time $t \in I$. It is worth to be highlighted that we let the variable T to run over $\mathscr{T}$, since we only deal with those bonds that do not expire within the first year.

Price curve dynamics. Throughout, we assume E to be represented by some space of continuous and real-valued functions defined on $\mathscr{T}$. We also assume the evaluation functional δ_T to be continuous and bounded on E, for any $T \in \mathscr{T}$. On the other hand, in order to capture the feature of any maturity specific-risk, it may be reasonable to let H be infinite dimensional. An instance of this setup is the one suggested by Carmona and Tehranchi [39], by considering the space introduced in the following definition.

Definition 6.6. Let $w : \mathscr{T} \to \mathbb{R}^+$ be a positive and increasing function. We write $\mathcal{H}_w$ to denote the space of absolutely continuous functions $x : \mathscr{T} \to \mathbb{R}$ with $x(s) \to 0$, as $s \to +\infty$, and such that

$$\int_{\mathscr{T}} x'(s)^2 w(s) ds < +\infty,$$

where x' stands for the weak derivative of x.

When endowed with the norm

$$\|x\|_{\mathcal{H}_w} \triangleq \left\{ x(1)^2 + \int_{\mathscr{T}} x'(s)^2 w(s) ds \right\}^{1/2}, \qquad \text{for any } x \in \mathcal{H}_w,$$

the space $\mathcal{H}_w$ turns out to be a Hilbert space, as reported in the Lemma below. In this respect, we recall that any Hilbert space is also a UMD Banach space with type 2.

Lemma 6.5. *Let $w : \mathscr{T} \to \mathbb{R}^+$ be a positive and increasing function such that*

$$\int_{\mathscr{T}} w(s)^{-1} ds < +\infty, \tag{6.47}$$

then $\mathcal{H}_w$ is a separable Hilbert space and the evaluation functional δ_T is continuous and bounded on $\mathcal{H}_w$, for any $T \in \mathscr{T}$.

Proof. See, e.g., Proposition 6.3 in [40]. $\square$

Here and in the sequel, we regard any element of E as the possible structure of the discounted price curve at a certain time. More precisely, we assume the risk-neutral dynamics of the discounted price curve to be governed by a certain adapted E-valued process $p = \{p_t : t \in I\}$. In this respect, let $\sigma = \{\sigma_t : t \in I\}$ be an adapted H-strongly measurable process such that $\sigma \in \mathbb{H}^{2,2}(L^2(I; \gamma(H, E)))$, and assume $p_0 \in \mathbb{H}^{1,2}(E)$ to be some strongly $\mathscr{G}_0^W$-measurable random variable. Then, we set

$$p_t = p_0 + \int_0^t \sigma_s dW_s, \qquad \text{for any } t \in I. \tag{6.48}$$

Besides, let $\delta_T \in E^*$ be the evaluation functional at $T \in \mathscr{T}$ and notice that

$$p_t(T) = \langle \delta_T, p_t \rangle_E, \qquad \text{for any } t \in I.$$

Thus, we may interpret δ_T as a portfolio composed by a zero coupon bond expiring at time $T \in \mathscr{T}$, and since the following equality a.s. holds true thanks to Theorem 5.2,

$$\left\langle \delta_T, \int_0^t \sigma_s dW_s \right\rangle_E = \int_0^t \langle \delta_T, \sigma_s \rangle_E dW_s, \qquad \text{for any } t \in I,$$

the identity (6.48) leads to the representation of the risk-neutral dynamics for the discounted price of the bond expiring at T given by

$$p_t(T) = p_0(T) + \int_0^t \sigma_s(T) dW_s, \qquad \text{for any } t \in I, \tag{6.49}$$

where for notation simplicity we set $\sigma_t(T) \triangleq \langle \delta_T, \sigma_s \rangle_E$.

Risk functional and portfolio duration. Let $f : I \times E \to \mathbb{R}$ be a function of class $\mathscr{C}_b^{1,2}$ and fix a P-set Φ relative to the process p given by the identity (6.48). We assume that for any $\phi \in \Phi$ there exists some $T \in \mathscr{T}$, such that

$$\phi_t = \alpha(T)\delta_T, \qquad \text{a.s., for any } t \in I, \tag{6.50}$$

where

$$\alpha(T) = p_0(T)^{-1} f(0, p_0). \tag{6.51}$$

As a direct result, we may write $\Phi = \mathscr{T}$ by identifying any process $\phi \in \Phi$ with $T \in \mathscr{T}$ such that the condition (6.50) is satisfied.

Proposition 6.4. *Let $f : I \times E \to \mathbb{R}$ be a function of class $\mathscr{C}_b^{1,2}$ and the process p be given by (6.48). Let Φ be a P-set relative to p, whose elements satisfy the identity (6.50). If f is a BS-function relative to p, then*

$$\mathcal{F}(T) = \mathbb{E}\left\{ \int_I \left\|(\nabla_2 f(t,p_t)\sigma_t - \alpha(T)\sigma_t(T)\right\|_H^2 (1-t)dt \right\}, \qquad \textit{for any } T \in \mathscr{T}. \quad (6.52)$$

Proof. The result follows directly form the statement (ii) in Theorem 6.1. Indeed, identify the process p with ξ, and notice that in the present case one has $b_t = 0$ a.s., for any $t \in I$. □

We may understand the variable $f(t, p_t)$ as the discounted value of a certain interest rate securities portfolio at time $t \in I$. We interpret any $\phi \in \Phi$ as the dynamics of a certain portfolio composed by a single bond relative to a fixed maturity and nominal given by the identity (6.51). Thus, we suppose the variable $\langle \phi_t, p_t \rangle_E$ to assess its risk-neutral discounted value at time $t \in I$. It is worth to be noted that the condition (6.51) is reasonable, since it guarantees that

$$f(0, p_0) = \langle \phi_0, p_0 \rangle_E,$$

and hence that any $\phi \in \Phi$ provides a perfect hedge relative to the portfolio represented by f at time $t = 0$. We may understand the E^*-valued variable $\nabla_2 f(t, \xi_t)$ as a notion of *duration* of the portfolio f at time $t \in I$, since it may represent the sensitivity of the portfolio to small changes in the structure of the price curve at time $t \in I$. In this respect, notice that if f is a BS-function relative to p and there exists $T^* \in \mathscr{T}$ such that

$$\nabla_2 f(t, p_t) = \alpha(T^*)\delta_{T^*} \qquad \text{a.s., for any } t \in I,$$

then the representation (6.52) gives $\mathcal{F}(T^*) = 0$, and hence T^* turns out to be $\mathcal{F}$-optimal. In this respect, when the financial exposure f is market valued, we may regard any $\mathcal{F}$-optimal $T^* \in \mathscr{T}$ as a notion of the related duration.

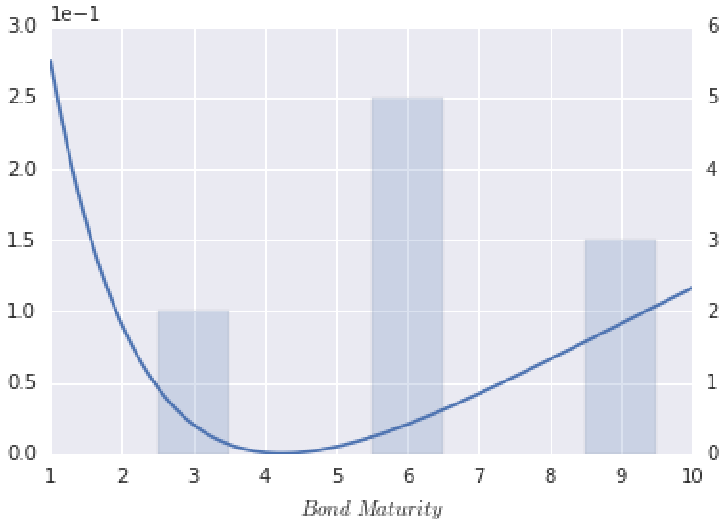

Figure 6.1: Visual representation of the functional (6.53) when considering the model (6.54) for the short-rate dynamics. More precisely, the line represents the functional (6.53) for any bond maturity $T \in (1, 10)$, the value of which is assessed by the left side vertical axis. The right side vertical axis assesses the nominals of the bonds within the fixed portfolio, that are represented by the vertical bars in the figure. The model (6.54) has been simulated with the parameters $a_1 = 0.12$, $a_2 = 0.1$, $\sigma_1 = 0.16$ and $\sigma_2 = 0.15$. For sake of simplicity, we considered the case $\varphi(t) = \varphi_1 + \varphi_2 t$, for any $t \in I$, where $\varphi_1 = 0.01$ and $\varphi_2 = 0.15$. Finally, we set $\varrho = -0.01$.

Numerical example. Figure 6.1 provides a visual representation of the functional (6.52), when considering a fixed portfolio constituted by three bonds with different nominals and maturities, and whose risk-neutral discounted value at time $t \in I$ is given by

$$f(t, p_t) = \sum_{k=1,2,3} \alpha_k p_t(T_k),$$

where $\alpha_k \in \mathbb{R}$ and $T_k \in \mathscr{T}$, for $k = 1, 2, 3$. In particular, the vertical bars stand for the nominal α_k of the bond related to any maturity T_k. The value of any α_k is assessed by the scale on the right side vertical axis of the figure. The line shows the behaviour of the functional,

$$\mathcal{F}(T) \triangleq \mathbb{E}\left\{ \int_I \left\| \sum_{k=1,2,3} \alpha_k \sigma_t(T_k) - \alpha(T)\sigma_t(T) \right\|_H^2 (1-t)dt \right\}, \qquad \text{for } T \in \mathscr{T}, \quad (6.53)$$

that is derived form (6.52) and whose values are assessed by the scale reported on the left side vertical axis of the graph.

We considered a correlated two-additive-factor Gaussian model governing the evolution of the short-rate process. More precisely, let (W_1, W_2) be a two-dimensional Wiener process with some instantaneous correlation $\varrho \in (-1, 1)$. The risk-neutral dynamics of the instantaneous-short-rate $r \triangleq \{r_t : t \in I\}$ is given by

$$r_t \triangleq \chi_{1,t} + \chi_{2,t} + \varphi(t), \qquad \text{for any } t \in I,$$

where for $i = 1, 2$ the process $\chi_i \triangleq \{\chi_{i,t} : t \in I\}$ is given by the following Vasicek-type model,

$$\chi_{i,t} = \chi_{i,0} - \int_0^t a_i \chi_{i,t} dt + \int_0^t \sigma_i dW_{i,t}, \qquad \text{for any } t \in I, \qquad (6.54)$$

for some given positive parameters a_i and σ_i jointly with the initial condition $\chi_{i,0} \in \mathbb{R}^+$, and where $t \in I \mapsto \varphi(t)$ is some deterministic function. We refer to Section 4.2. in [25] for all the details.

This model is analytically tractable enough to write down an explicit formula for the discounted price curve in terms of the short-rate factors (6.54), and hence to determine an analytical expression for the process σ by considering standard calculus techniques. For any $T \in \mathscr{T}$, the value $\mathcal{F}(T)$ has been obtained as a combination of the Monte Carlo simulation of the factors (6.54) jointly with the the discretization of the integral related to the time variable $t \in I$. All the details of the simulation are collected in the caption of Figure 6.1.

CHAPTER 7

Optimal Model Points Portfolio

European insurance companies are required to assess the value of their portfolios as well as to carry on the sensitivity analysis aimed at demonstrating the compliance of their models, by considering the cash flow projections on a policy-by-policy approach. Besides, they are allowed to compute these projections by replacing any homogeneous group of policies with some suitable representative contracts, usually known as the related model points. All this, in order to speed up this process, that is usually carried out on a daily basis, since the complexity of the entire portfolio may lead to long computational times. This procedure is permitted under suitable conditions in such a way that the inherent risk structure of the original portfolio is not misrepresented. We refer to [68] for further details.

In this chapter, we assess the problem of determining an optimal model points portfolio related to some fixed policies portfolio by applying the results discussed in Chapter 6. In particular, we characterize the models points associated to a given life insurance portfolio as the set of policies that minimizes a certain risk functional that we introduce in a similar manner as the one presented in §6.1. As a particular instance, we show how these arguments apply when considering a portfolio of whole life insurance policies.

This chapter is based on the original work:

[79] Enrico Ferri. *Optimal Model Points Portfolio in Life Insurance.* arXiv:1808.00866, 2018.

7.1 LIFE INSURANCE MODEL POINTS PORTFOLIO

Life insurance contracts usually provide either a stream of cash flow during the lifetime of the policyholder or a unique lump sum benefit that is paid upon the death of the insured and under certain conditions. According to the clauses of the contract, other events like the appearance of a terminal illness may also lead to an early payment. The contract is alive when the policyholder pays a specific premium, that depending upon the case may be either regular or as one initial lump sum.

We do not consider the specific characteristics of the contract and we assume to deal with some idealized life insurance policy that is unaffected by credit risk, i.e. the insurance company always guarantees the entire benefit that is provided for in the contract. Furthermore, we do not analyse the revenues received by the insurance company and thus we do not take into account the premium payments associated the contract such as any further expenses that are responsibility of the client.

Let $\mathscr{X}$ be some set in which any element $x \in \mathscr{X}$ represents a contract. To fix the ideas, any $x \in \mathscr{X}$ may collect typical characteristics such as age and gender of the policy owner, cancellation option, etc. In other terms, any element $x \in \mathscr{X}$ identifies a class of policies that is labelled by considering certain suitable characteristics. We make the convenient assumption that there exists a market value of the contract relative to any $x \in \mathscr{X}$.

Model points portfolios. In the sequel, we fix a UMD Banach space U that is represented by some space of real valued functions defined on $\mathscr{X}$. We write U^* to denote the topological dual of U. The duality pairing between U and U^* is denoted by $\langle \cdot, \cdot \rangle_U$. Moreover, we assume the evaluation functional δ_x to be continuous and bounded on U, for any $x \in \mathscr{X}$. Thus, we write $\mathscr{X}^*$ to denote the subsets of U^* defined by setting

$$\mathscr{X}^* \triangleq \overline{\operatorname{span}\{\delta_x : x \in \mathscr{X}\}}, \tag{7.1}$$

where the closure in (7.1) is understood with respect to the topology of U^*. Moreover, we fix a subset $\mathscr{Y}^* \subseteq \mathscr{X}^*$.

Definition 7.1. Let $v \in \mathscr{X}^*$ and consider a U-valued process $z \triangleq \{z_t : t \in I\}$. For any $w \in \mathscr{Y}^*$, the process $V_t(w) \triangleq \{V_t(w) : t \in I\}$ defined by setting

$$V_t(w) = \langle v, z_t \rangle_U - \langle w, z_t \rangle_U, \qquad \text{for any } t \in I, \tag{7.2}$$

is said to be the *discrepancy process* between v and w relative to z.

If not otherwise specified, where we fix a U-valued process $z \triangleq \{z_t : t \in I\}$, for any $v \in \mathscr{X}^*$ and $w \in \mathscr{Y}^*$, we always write $V(w)$ to denote the discrepancy process between v and w relative to z. For our application, we regard any process $z = \{z_t : t \in I\}$ as a dynamics for the discounted value of some specific life insurance contract. More precisely, we understand $z_t(x)$ as the risk-neutral discounted value of the contract at time $t \in I$, when it is referred to $x \in \mathscr{X}$.

Besides, similarly to the framework of Section 6.4, it is worth to be noted that

$$z_t(x) = \langle \delta_x, z_t \rangle_U, \qquad \text{for any } t \in I \text{ and } x \in \mathscr{X}.$$

Hence, we may regard δ_x as a portfolio composed by one policy related to $x \in \mathscr{X}$ and any $v \in \mathscr{X}^*$ as a portfolio composed by different policies. On the other hand, we regard any $w \in \mathscr{Y}^*$ as a specific model points portfolio.

Lemma 7.1. *Let $z \triangleq \{z_t : t \in I\}$ be a U-valued process such that $\sup_{t\in I} \|z_t\|^2_{L^2(\Omega;U)} < \infty$ and fix $v \in \mathscr{X}^*$. Then, $V_t(w) \in L^2(\Omega)$, for any $t \in I$ and $w \in \mathscr{Y}^*$, with*

$$\sup_{t\in I} \|V_t(w)\|^2_{L^2(\Omega)} < \infty.$$

Proof. Fix $w \in \mathscr{Y}^*$ and notice that

$$\sup_{t\in I} \|V_t(w)\|^2_{L^2(\Omega)} = \sup_{t\in I} \|\langle v - w, z_t\rangle_U\|^2_{L^2(\Omega)} \leq \|v - m\|^2_{U^*} \sup_{t\in I} \|z_t\|^2_{L^2(\Omega;U)}.$$

□

In view of our applications, we recast Definitions 6.4 and 6.5 as follows.

DEFINITION 7.2. Let $z \triangleq \{z_t : t \in I\}$ be an U-valued process such that

$$\sup_{t\in I} \|z_t\|^2_{L^2(\Omega;U)} < \infty. \tag{7.3}$$

and fix $v \in \mathscr{X}^*$. We refer to the functional $\mathcal{V} : \mathscr{Y}^* \to \mathbb{R}$ defined by

$$\mathcal{V}(w) \triangleq \int_I \mathbb{E}\{|V_t(w) - \mathbb{E}V_t(w)|^2\}dt, \qquad \text{for any } w \in \mathscr{Y}^*, \tag{7.4}$$

as the *model points risk functional* relative to z induced by v over the set $\mathscr{Y}^*$. Moreover, we call *optimal model points portfolio* any $\mathcal{V}$-optimal element $w^* \in \mathscr{Y}^*$.

If not otherwise specified, where we fix a process $z \triangleq \{z_t : t \in I\}$ satisfying the condition (7.3) and a portfolio $v \in \mathscr{X}^*$ is given, we always write $\mathcal{V}$ to denote the model points risk functional relative to z induced by v over $\mathscr{Y}^*$.

Model points risk functional representation. Let E and H be the spaces as considered in Section 6.4 and thus assume the process $p = \{p_t : t \in I\}$ defined by the identity (6.48) to model the risk-neutral dynamics of the discounted price curve. Since the value of any policy is required to be estimated by considering the probability weighted average of the related future cash flow, it is reasonable to write

$$z_t = \zeta(t, p_t) \qquad \text{for any } t \in I,$$

for some function $\zeta : I \times E \to U$. We discuss this approach later on.

Lemma 7.2. *Fix a function $\zeta : I \times E \to U$ and denote by p the process (6.48). If ζ is of class $\mathscr{C}^{1,2}_b$, then $\zeta(t, p_t) \in L^2(\Omega; U)$ is well defined, for any $t \in I$, with*

$$\sup_{t\in I} \|\zeta(t, p_t)\|^2_{L^2(\Omega;U)} < \infty.$$

Proof. First, notice that since $\zeta(\cdot, 0)$ is assumed to be continuous on I, we have that

$$\|\zeta(\cdot, 0)\|_\infty \triangleq \sup_{t\in I} \|\zeta(t, 0)\|_U < \infty.$$

Besides, since ζ is assumed to be of class $\mathscr{C}_b^{1,2}$, for any $t \in I$, the following inequalities hold true a.s.,

$$\|\zeta(t, p_t)\|_U^2 \leq \|\nabla_2\zeta\|_\infty^2 \|p_t\|_E^2 + |\zeta(t, 0)|^2 < \infty,$$

and hence

$$\sup_{t\in I} \|\zeta(t, p_t)\|_U^2 \leq \|\nabla_2\zeta\|_\infty^2 \sup_{t\in I} \|p_t\|_E^2 + \|\zeta(\cdot, 0)\|_\infty^2 < \infty,$$

since $\sup_{t\in I} \|p_t\|_E^2 < \infty$ due to Lemma 5.8, applied to the process p. □

The following result characterizes the model points risk functional relative to z, when it is given by the identity (7.5), induced by a portfolio v over $\mathscr{Y}^*$.

Proposition 7.1. *Fix $v \in \mathscr{X}^*$ and let $\zeta : I \times E \to U$ be a function of class $\mathscr{C}_b^{1,2}$. If ζ is a BL-function relative to p, where the process p is given by (6.48), and we set $z_t = \zeta(t, p_t)$, for any $t \in I$, then*

$$\mathcal{V}(w) = \mathbb{E}\left\{ \int_I \|\langle v - w, \nabla_2\zeta(t, p_t)\sigma_t\rangle_U\|_H^2 (1-t)dt \right\}, \qquad \textit{for any } w \in \mathscr{Y}^*. \tag{7.5}$$

It is worth to be noted that that, since the process σ takes values in $\gamma(H, E)$, Lemma 5.1 gives that the process $\nabla_2\zeta(t, p_t)\sigma_t$, for $t \in I$, takes values in $\gamma(H, U)$. Thus, in the representation (7.5) we regard $\langle\cdot,\cdot\rangle_U$ as the H-valued pairing between $\gamma(H, U)$ and U^*.

Proof of Proposition 7.1. First, notice that $\sup_{t\in I} \|z_t\|_{L^2(\Omega;U)}^2 < \infty$ thanks to Lemma 7.2, since the function $\zeta : I \times E \to U$ is assumed to be of class $\mathscr{C}_b^{1,2}$. On the other hand, notice that since ζ is assumed to be a BL-function relative to p, and the process p is given by (6.48), Lemma 6.1 gives that a.s.

$$\zeta(t, p_t) = \zeta(0, p_0) + \int_0^t \nabla_2\zeta(s, p_s)\sigma_s dW_s, \quad \text{for any } t \in I. \tag{7.6}$$

Let Φ to be a P-set relative to z defined in such a way that for any $\phi \in \Phi$ there exists $w \in \mathscr{Y}^*$ with $\phi_t = w$, a.s., for any $t \in I$. Then, the result follows from the statement (ii) in Theorem 6.1 by setting $\xi_t = \zeta(t, p_t)$ and $f(t, u) = \langle v, u\rangle_U$, for any $u \in U$. In this respect, notice that $b_t = 0$ a.s., for any $t \in I$, and f appears to be a BS-function relative to z. $\square$

7.2 WHOLE LIFE INSURANCE

In this section we discuss the problem of the model points selection when dealing with a portfolio of whole life insurance policies, i.e. a contract that provides for a one unit benefit on the death of the policy owner. In view of this application, we assume $\mathscr{X}$ to coincide with some closed interval of $\mathbb{R}^+$ and we regard any $x \in \mathscr{X}$ as the age of the policy owner at time $t = 0$.

Moreover, here and in the sequel of this section we define $U \triangleq \mathbb{W}^{1,2}(\mathscr{X})$ to be the space of absolutely continuous functions $u : \mathscr{X} \to \mathbb{R}$ such that $\|u'\|_{L^2(\mathscr{X})} < \infty$, where we denoted by u' the weak derivative of u. In this respect, recall that U is a Hilbert space, when endowed with the norm

$$\|u\|_U \triangleq \{\|u\|^2_{L^2(\mathscr{X})} + \|u'\|^2_{L^2(\mathscr{X})}\}^{1/2}, \qquad \text{for any } u \in U,$$

such that the evaluation functional $\delta_x(u) \triangleq u(x)$, for $u \in U$, is bounded for any $x \in \mathscr{X}$.

Force of mortality and whole life policy process. Here and in the sequel, we write $\mu(s, x + s)$ to denote the force of mortality relative to $x \in \mathscr{X}$ at any time $s \geq 0$, i.e. the instantaneous rate of mortality at time s and relative to an individual that is x-aged at time $t = 0$. We make the convenient assumption that the force of mortality $\mu(s, x + s)$ is given at any $s \geq 0$ and for any $x \in \mathscr{X}$. Moreover, we do not consider those deaths that occur during the first year by setting

$$\mu(s, x + s) = 0, \qquad \text{for any } x \in \mathscr{X} \text{ and for } s \in I, \tag{7.7}$$

and hence defining the survival index as

$$S(x, T) \triangleq \exp\left\{ -\int_1^T \mu(s, x + s)ds \right\}, \qquad \text{for any } x \in \mathscr{X} \text{ and } T \in \mathscr{T}. \tag{7.8}$$

We regard $S(x,T)$ as the proportion of individuals that are x-aged at time $t=0$ and that survive to age $x+T$.

Note that condition (7.7) results convenient for our application, since it implies that the policy portfolio does not change during the time interval I due to the death of the policy owners. Such an assumption is acceptable, since the events occurring within the first year only cause a minimal impact on the performance of the overall portfolio.

The discounted value of a whole life insurance policy relative to $x \in \mathscr{X}$ at time $t \in I$ may be written as

$$z_t(x) = \int_{\mathscr{T}} S(x,T)\mu(T,x+T)p_t(T)dT. \tag{7.9}$$

The following Lemma tell proves under mild conditions that the identity (7.9) provides a well defined U-valued process.

Lemma 7.3. *Let $w : \mathscr{T} \to \mathbb{R}^+$ be an increasing function satisfying (6.47) and set $E = \mathcal{H}_w$ according to Definition 6.6. Moreover, assume the function $x \in \mathscr{X} \mapsto \mu(s, x+s)$ to be continuously differentiable, for any $s \in \mathscr{T}$. If $w \geq 1$ everywhere on $\mathscr{T}$ and the following condition holds*

$$\sup_{T\in\mathscr{T}} w(T)^{-1} \int_{\mathscr{X}} |\partial_x S(x,T)|^2 dx < \infty, \tag{7.10}$$

then the process $z \triangleq \{z_t : t \in I\}$ given by the identity (7.9) is U-valued, with

$$\sup_{t\in I} \|z_t\|^2_{L^2(\Omega;U)} < \infty. \tag{7.11}$$

Proof. First, notice that since $p_t(T) \to 0$ a.s., as $T \to +\infty$, for any $t \in I$, and

$$S(x,T)\mu(T,x+T) = -\partial_T S(x,T), \qquad \text{for any } x \in \mathscr{X} \text{ and } T \in \mathscr{T},$$

then the identity (7.9) may be recast by invoking integration by parts arguments as follows,

$$z_t(x) = -S(x,1)p_t(1) + \int_{\mathscr{T}} S(x,T)p_t'(T)dT, \qquad \text{for any } t \in I \text{ and } x \in \mathscr{X}.$$

As a direct consequence, for any fixed $t \in I$, we obtain

$$\begin{aligned}\int_{\mathscr{X}} z_t(x)^2 dx &\leq \int_{\mathscr{X}} S(x,1)^2 p_t(1)^2 dx + \int_{\mathscr{X}} \int_{\mathscr{T}} S(x,T)^2 p_t'(T)^2 dT dx \\ &\overset{\text{(i)}}{\leq} p_t(1)^2 + m(\mathscr{X}) \int_{\mathscr{T}} p_t'(T)^2 dT, \end{aligned} \tag{7.12}$$

where the inequality (i) holds true since $S(x,T) \leq 1$, for any $x \in \mathscr{X}$ and $T \in \mathscr{T}$, and the function $x \in \mathscr{X} \mapsto \mu(s, x+s)$ is continuous, for any $s \in \mathscr{T}$. In (7.12) we denoted by $m(\mathscr{X})$ the Lebesgue measure of the interval $\mathscr{X}$.

Next, Lemma 5.8 applied to the process p implies that

$$\sup_{t\in I} \mathbb{E}\|p_t\|_E^2 = \sup_{t\in I} \mathbb{E}\left\{ \int_{\mathscr{T}} p'(T)^2 w(T) dT \right\} < \infty. \tag{7.13}$$

As a consequence, Cauchy-Schwartz inequality jointly with Lemma 6.5 lead to

$$\sup_{t\in I} \mathbb{E}\{p_t(1)^2\} = \sup_{t\in I} \mathbb{E}|\langle \delta_1, p_t\rangle_E|^2 \leq \|\delta_1\|_{E^*}^2 \sup_{t\in I} \mathbb{E}\|p_t\|_E^2 < \infty, \tag{7.14}$$

and Hölder's inequality jointly with the condition $w \geq 1$, everywhere on $\mathscr{T}$, implies that

$$\sup_{t\in I} \mathbb{E}\left\{ \int_{\mathscr{T}} p_t'(T)^2 dT \right\} \leq \sup_{t\in I} \mathbb{E}\left\{ \int_{\mathscr{T}} p_t'(T)^2 w(T) dT \right\} \sup_{T\in\mathscr{T}} w(T)^{-1} < \infty. \tag{7.15}$$

Next, the inequalities (7.14) and (7.15) combined to (7.12) give

$$\sup_{t\in I} \mathbb{E}\|z_t\|_{L^2(\mathscr{X})}^2 = \sup_{t\in I} \mathbb{E}\left\{ \int_{\mathscr{X}} z_t(x)^2 dx \right\} < \infty.$$

On the other hand, since the function $x \in \mathscr{X} \mapsto \mu(s, x+s)$ is assumed to be continuously differentiable, then

$$T \in \mathscr{T} \mapsto \int_{\mathscr{X}} |\partial_x S(x,T)|^2 dx,$$

is well defined everywhere on $\mathscr{T}$. Furthermore,

$$z_t'(x) = -\partial_x S(x,1) p_t(1) + \int_{\mathscr{T}} \partial_x S(x,T) p_t'(T) dT, \qquad \text{for any } t \in I \text{ and } x \in \mathscr{X}, \tag{7.16}$$

and hence, for any fixed $t \in I$, we get

$$\int_{\mathscr{X}} z_t'(x)^2 dx \leq p_t(1)^2 \int_{\mathscr{X}} |\partial_x S(x,1)|^2 dx + \int_{\mathscr{T}} \left\{ \int_{\mathscr{X}} |\partial_x S(x,T)|^2 dx \right\} p_t'(T)^2 dT. \quad (7.17)$$

Next, since

$$\sup_{t \in I} \mathbb{E} \left\{ \int_{\mathscr{T}} \left\{ \int_{\mathscr{X}} |\partial_x S(x,T)|^2 dx \right\} p_t'(T)^2 dT \right\} \leq \sup_{t \in I} \mathbb{E} \left\{ \int_{\mathscr{T}} p_t'(T)^2 w(T) dT \right\} \sup_{T \in \mathscr{T}} w(T)^{-1} \int_{\mathscr{X}} |\partial_x S(x,T)|^2 dx < \infty,$$

thanks to Hölder's inequality again combined with the assumption (7.10) and the condition (7.13), according to (7.14) we get that (7.17) leads to

$$\sup_{t \in I} \mathbb{E} \| z_t' \|^2_{L^2(\mathscr{X})} = \sup_{t \in I} \mathbb{E} \left\{ \int_{\mathscr{X}} z_t'(x)^2 dx \right\} < \infty.$$

□

Whole life model points portfolio. In §6.4 we represented an interest rate securities portfolio by considering a portfolio composed by a single zero coupon bond. In a similar manner, throughout we consider the case in which a portfolio of whole life insurance policies owned by individual with different ages and time $t = 0$ is represented by considering a certain amount of whole life policies relative to the same age.

With this purpose in mind, we fix $x_1, ..., x_K \in \mathscr{X}$ and $\alpha_1, ..., \alpha_K \in \mathbb{R}^+$ and we consider the portfolio $v \in \mathscr{X}^*$ defined by setting

$$v \triangleq \sum_{k=1}^{K} \alpha_k \delta_{x_k}. \quad (7.18)$$

On the other hand, we assume that for any $w \in \mathscr{Y}^*$ there exists $x \in \mathscr{X}$ such that the following representation holds

$$w = \alpha(x) \delta_x, \quad (7.19)$$

for some $x \in \mathscr{X}$, where

$$\alpha(x) \triangleq z_0(x)^{-1}\langle v, z_0\rangle_U. \tag{7.20}$$

Therefore, we may write $\mathscr{Y}^* = \mathscr{X}$, by identifying any $w \in \mathscr{Y}^*$ with the element $x \in \mathscr{X}$ that satisfies the representation (7.19). It is worth to be noted that the condition (7.20) is reasonable, since it guarantees that

$$\langle v, z_0\rangle_U = \langle w, z_0\rangle_U,$$

and hence that any model points portfolio $w \in \mathscr{Y}^*$ admits the same discounted value as $v \in \mathscr{X}^*$, at time $t = 0$.

Proposition 7.2. *Let $v \in \mathscr{X}^*$ as defined by the identity (7.18) and $\mathscr{Y}^*$ such that the representation (7.19) holds true for any $w \in \mathscr{Y}^*$. Furthermore, let $z = \{z_t : t \in I\}$ be the process defined by the identity (7.9). Then, we obtain*

$$\mathcal{V}(x) = \mathbb{E}\left\{ \int_I \left\| \int_{\mathscr{T}} \left(\sum_{k=1}^{K} \alpha_k \kappa(x_k, T) - \alpha(x)\kappa(x, T)\right) \sigma_t(T) dT \right\|_H^2 (1-t)dt \right\}, \quad \textit{for any } x \in \mathscr{X}, \tag{7.21}$$

where

$$\kappa(x, T) \triangleq S(x, T)\mu(T, x+T), \qquad \textit{for any } x \in \mathscr{X} \textit{ and } T \in \mathscr{T}.$$

Proof. First, we introduce the functional $Z \in \mathcal{L}(E, U)$ defined

$$Z(q) = \int_{\mathscr{T}} \kappa(\cdot, T)q(T)dT, \qquad \text{for any } q \in E.$$

Hence, when considering the function $\zeta : I \times E \to U$ given by the identity

$$\zeta(t, q) \triangleq Z(q), \qquad \text{for any } t \in I \text{ and } q \in E, \tag{7.22}$$

we obtain that ζ is of class $\mathscr{C}_b^{1,2}$ and it turns out to be a BL-function relative to p. Moreover, note that the process $z = \{z_t : t \in I\}$ defined by the identity (7.9) is recovered by setting $z_t = Z(p_t)$, for any $t \in I$. In this respect, it is worth to be highlighted that the identification (7.22) is allowed since the survival index (7.8) does not depend on $t \in I$, that is the case when imposing the condition (7.7). Thus, when writing ∇Z to denote the operator $\nabla_2 \zeta$, Proposition 7.1 applies and leads to

$$\mathcal{V}(x) = \mathbb{E}\left\{ \int_I \left\| \sum_{k=1}^{K} \alpha_k \nabla Z(p_t)\sigma_t(x_k) - \alpha(x)\nabla Z(p_t)\sigma_t(x) \right\|_H^2 (1-t)dt \right\}, \qquad \text{for any } x \in \mathscr{X}, \tag{7.23}$$

where, for notation simplicity, in the identity (7.23) we set

$$\nabla Z(p_t)\sigma_t(x) \triangleq \langle \delta_x, \nabla Z(p_t)\sigma_t \rangle_U, \qquad \text{for any } t \in I \text{ and } x \in \mathscr{X}.$$

On the other hand, by a direct computation, one has that a.s.

$$\nabla Z(p_t)\sigma_t(x) = \int_{\mathscr{T}} \kappa(x,T)\sigma_t(T)dT, \qquad \text{for any } x \in \mathscr{X},$$

and hence, jointly with the identity (7.23) we obtain the representation (7.21). $\square$

Numerical example. Figure 7.1 provides a visual representation of the functional (7.21). In particular, each bar represents the amount α_k associated to the age x_k, for $k = 1, ..., K$, and it is assessed by the right side vertical axis of the figure. On the other hand, the functional $\mathcal{V}(x)$, varying $x \in \mathscr{X}$, is represented by the line and its value is assessed by the scale on the left side vertical axis of the figure.
The process p has been simulated by considering the same model governing the evolution of the instantaneous-short-rate as in Section 6.4. The survival index (7.8) has been derived by considering a Gompertz-type law modelling the force of mortality [94], defined by

$$\mu(s, x+s) = a(s)\exp\{(x+s)b(s)\}, \qquad \text{for any } x \in \mathscr{X} \text{ and } s \in \mathscr{T}. \tag{7.24}$$

where $a(s)$ and $b(s)$, varying $s \in \mathscr{T}$, are positive functions. All the details of the simulation are collected in the caption of Figure 7.1.

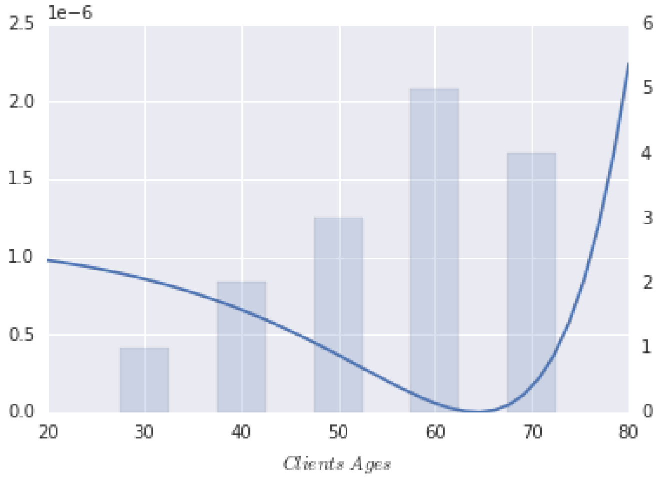

Figure 7.1: Visual representation of the functional (7.5), where $K = 5$, with $x_1 = 30$, $x_2 = 40$,...,$x_K = 70$. In particular, the values α_k, for any $k = 1, ..., K$, are represented by the vertical bar and they are assessed by the right side vertical axis. The value $\mathcal{V}(x)$, for $20 \leq x \leq 80$, is represented by the blue line and it is assessed by the left side vertical axis. The process (6.54) has been considered to model the short rate dynamics with the same parameters as described in the caption of Figure 6.1. For simplicity, the mortality force (7.24) has been computed by setting $a(s) = 0.0003$ and $b(s) = 0.06$, for any $s \geq 1$.

CHAPTER 8

LIBOR Market Model in Term Insurance

In Chapter 7, for a given a set of model points portfolios $\mathcal{W}$ and a given model points risk functional $\mathcal{V} : \mathcal{W} \to \mathbb{R}$ induced by the policy portfolio v over $\mathcal{W}$, any element $w^* \in \mathcal{W}$ which solves the minimization problem

$$w^* = \mathrm{argmin}_{w \in \mathcal{W}} \mathcal{V}(w), \tag{8.1}$$

has been called $\mathcal{V}$-optimal. In this respect, recall that any $\mathcal{V}$-optimal portfolio $w^* \in \mathcal{W}$ may be understood as the best representation of the inherent risk of v, among all the portfolios $w \in \mathcal{W}$.

From the numerical point of view, in most practical situation such a minimization results to be computationally demanding, since it generally leads to a global optimization problem in high dimensions. On the other hand, the numerical simulation of the interest rates dynamics may involve high computational costs, which is to be addressed by using proper computer architectures and efficient simulation techniques.

In this chapter, we study the performance of certain numerical methods aimed at adressing the minimization problem (8.1), when considering a portfolio of term insurance policies and a LIBOR Market Model governing the dynamics of the forwards rates. Since one of the main features is represented by the forward rate dynamics, throughout we consider the standard notation for the LIBOR Market model, that is different from which we previously used.

This chapter is based on the original work [76].

8.1 BOND DYNAMICS IN A LIBOR MARKET MODEL

In this section, we discuss the risk-free dynamics of the discounted bond price, when considering the LIBOR Market Model governing the time evolution of the forward rates.

Preliminaries. Here and in the sequel, we fix a finite set $\mathscr{T} \triangleq \{T_0, T_1, ..., T_N\}$ such that $T_0 = 1$ and $T_0 < ... < T_N$. We shall regard any T_n for any $n = 0, ..., N$ as a specific maturity time and we shall write $\tau_n \triangleq T_n - T_{n-1}$. Moreover, $I \triangleq (0, 1)$ denotes the unit interval on the real line and it corresponds to the period of one year.

We shall suppose that there exists a market valued bond maturing at any future time $T_n \in \mathscr{T}$ and we shall write $p_n(t)$ to denote its risk-neutral discounted price at time $t \in I$. Moreover, we shall denote by $F_n(t)$ the value at time $t \in I$ of the (simply-compounded) LIBOR forward rate associated to the period $(T_{n-1}, T_n]$. In this respect, recall that $F_n(t)$ is defined in such a way that the following condition is met,

$$p_n(t)(1 + F_n(t)\tau_n) = p_{n-1}(t).$$

Hence, we can write

$$p_n(t) = p_0(t) \prod_{k=1}^{n} \frac{1}{1 + F_k(t)\tau_k}, \qquad \text{for any } t \in I. \tag{8.2}$$

In this connection, it is worth to be highlighted that since $t < T_n$, for any $t \in I$ and any $T_n \in \mathscr{T}$, the quantity $p_n(t)$ is always well defined.

LIBOR Market Model. Here and in the sequel, we consider a complete probability space $(\Omega, \mathscr{F}, \mathbb{P})$ and we assume any stochastic process to be defined on it. Moreover, we set $H = \mathbb{R}^N$ and for any $h \in H$, we denote by $h_1, ..., h_N$ its components with respect to a given orthonormal basis. Hence, let $\| \cdot \|_H$ denote the norm on it defined by

$$\|h\|_H \triangleq \left\{ \sum_{n,k} \varrho_{nk} h_n h_k \right\}^{1/2}, \qquad \text{for any } h \in H,$$

where $\varrho_{nk} = e^{-\delta(n-k)}$ for some constant $\delta \in \mathbb{R}^+$.

Here and in the sequel, we fix a H-cylindrical process $W \triangleq \{W(t) : t \in I\}$. According to similar arguments to those ones presented in Section 6.3, the process W may be regarded as a correlated N-dimensional Weiner process $W(t) = (W_1(t), ..., W_N(t))$, for $t \in I$, such that

$$\mathbb{E}\{W_n(t)W_k(t)\} = \varrho_{nk}.$$

As in the previous chapters, we say a process to be adapted if it is adapted with respect to the augmented filtration $\mathscr{G}^W \triangleq \{\mathscr{G}_t^W : t \in I\}$ generated by W.

Throughout, let $p_0(t)$, for $t \in I$, be an adapted process and $\mathbb{P}$ the forward measure associated to the maturity $T_0 \in \mathscr{T}$, i.e. such that for any adapted process $X \triangleq \{X(t), t \in I\}$, the associated discounted process $\tilde{X}(t) \triangleq \{\tilde{X}(t) : t \in I\}$ given by

$$\tilde{X}(t) = p_0(t)^{-1} X(t), \qquad \text{for any } t \in I,$$

turns out to be a $\mathscr{G}^W$-martingale. According to the LIBOR Market Model (see, e.g., Section 6 in [25]), the forward rate $F_n(t)$, for $t \in I$, is thus given by the dynamics

$$dF_n(t) = \mu_n(t)dt + \sigma_n(t)F_n(t)dW_n(t), \tag{8.3}$$

jointly with some given initial condition $F_n(0)$. Here and in the sequel, $t \in I \mapsto \sigma_n(t)$ are deterministic functions and $\mu_n(t)$, for $t \in I$, are completely determined by the following identity,

$$\mu_n(t) = \sigma_n(t)F_n(t) \sum_{k=1}^{n} \frac{\varrho_{nk}\tau_k\sigma_k(t)F_k(t)}{1 + F_k(t)\tau_k}.$$

On the other hand, the identity (8.3), for $n = 1, ..., N$, may be understood as the dynamics of the components of the N-dimensional process $F(t) = (F_1(t), ..., F_N(t))$,

$$dF(t) = \mu(t)dt + \Sigma(t)dW(t),$$

jointly with the initial condition $F(0) = (F_1(0), ..., F_N(0))$, where, for any fixed $t \in I$, we set $\mu(t) = (\mu_1(t), ..., \mu_N(t))$ and $\Sigma(t)$ denotes the matrix whose components are given by

$$\Sigma_{nk}(t) = \sigma_n(t) F_n(t) \delta_{nk}.$$

Throughout, for notation simplicity we shall write $\Sigma_n(t)$ to denote the nth row of $\Sigma(t)$.

Bond price dynamics. Here and in the sequel, the discounted price at $t \in I$ associated to the bond expiring at the tenor date $T_n \in \mathscr{T}$ is denoted by

$$\tilde{p}_n(t) \triangleq \frac{p_n(t)}{p_0(t)}, \qquad \text{for any } t \in I. \tag{8.4}$$

According to the present setup, the process $\tilde{p}_n(t)$, for $t \in I$, turns out to be a $\mathscr{G}^W$-martingale. Moreover, its dynamics is given by the following result.

Lemma 8.1. *For any $n = 1, ..., N$, the discounted bond price process $\tilde{p}_n(t)$ admits the dynamics*

$$d\tilde{p}_n(t) = -\varepsilon_n(t) \tilde{p}_n(t) dW(t), \tag{8.5}$$

where

$$\varepsilon_n(t) \triangleq \sum_{k=1}^{n} \frac{\tau_k}{1 + F_k(t)\tau_k} \Sigma_k(t). \tag{8.6}$$

Proof. Fix $n = 1, ..., N$ and notice that according to the identity (8.2) the discounted price $\tilde{p}_n(t)$ at time $t \in I$ given in (8.4) is such that

$$\ln \tilde{p}_n(t) = -\sum_{k=1}^{n} \ln(1 + F_k(t)\tau_k).$$

For any diffusion process $X(t)$, for $t \in I$, driven by the Wiener process W, let $\mathrm{DC}X(t)$ denote its vector diffusion coefficient. In particular, the representation (8.4) implies that

$$\mathrm{DC}F_n(t) = \Sigma_n(t), \tag{8.7}$$

for any $t \in I$. Thus, we obtain

$$\begin{aligned}
\mathrm{DC} \ln \tilde{p}_n(t) &= -\sum_{k=1}^{n} \mathrm{DC} \ln(1 + \tau_k F_k(t)) \\
&= -\sum_{k=1}^{n} \frac{\tau_k}{1 + \tau_k F_k(t)} \mathrm{DC} F_k(t) \\
&= -\sum_{k=1}^{n} \frac{\tau_k}{1 + \tau_k F_k(t)} \Sigma_k(t). \qquad (8.8)
\end{aligned}$$

In order to complete the proof, notice that

$$\mathrm{DC}\tilde{p}_n(t) = p_n(t) \mathrm{DC} \ln \tilde{p}_n(t), \qquad \text{for any } t \in I.$$

□

8.2 TERM INSURANCE POLICIES

In this section, we discuss the problem of the model points selection when dealing with a portfolio of term insurance policies, i.e. those contracts that pay a lump sum benefit on the death of the policy owner, provided that it occurs until a specific term that is defined in the contract. For the sake of simplicity, we assume that the benefit related to each policy is always represented by a unit amount of a certain currency.

We assume to deal with policies that are unaffected by credit risk, i.e. the insurance company always guarantees the entire benefit that is provided for in the contract. On the other hand, we do not analyse the revenues received by the insurance company and thus we do not take into account the premiums stream of the contract nor any further expense that is responsibility of the client.

Term insurance portfolios. We assume the generic term insurance policy within a given portfolio to be labelled by both the age of the policy owner at time $t = 0$ and the term date of the contract. For this purpose, given two finite index sets $\mathcal{I}$ and $\mathcal{J}$, let $\mathcal{X} = \{x_i : i \in \mathcal{I}\}$ and $\mathcal{Y} = \{y_j : j \in \mathcal{J}\}$ be finite sequences of real values such that $x_i \geq 0$, for any $i \in \mathcal{I}$ and $y_j \geq 1$, for any and $j \in \mathcal{J}$. Here and in the sequel, any couple (x_i, y_j) uniquely defines the family of policies related to the class of individuals labelled by $x_i \in \mathcal{X}$ and with maturity $y_j \in \mathcal{Y}$.

Hereafter, we shall write $\mu(s, x_i + s)$ to denote the force of mortality at time $s \geq 0$ related to the class of individuals labelled by $x_i \in \mathcal{X}$. Moreover, we set

$$\mu(s, x_i + s) \triangleq a(s) \exp{(x_i + s) b(s)}, \qquad \text{for any } s \geq 0 \text{ and } x_i \in \mathcal{X}, \tag{8.9}$$

where $a(s)$ and $b(s)$, for $s \geq 0$, are given functions, which are considered to be deterministic observables. We make the convenient assumption that $a(s) = 0$, for any $s \in I$, and hence

$$\mu(s, x_i + s) = 0, \qquad \text{for any } i \in \mathcal{I} \text{ and any } s \in I, \tag{8.10}$$

which guarantees that any portfolio of term insurance policies does not change within the time interval I due to the death of the policy holders. Recall that this is an acceptable hypothesis, since the events occurring within the first year only cause a minimal impact on the performance of the overall portfolio.

According to this setup, the survival index is thus given by

$$S(x_i, T_n) = \exp\left\{ -\int_1^{T_n} \mu(s, x_i{+}s) ds \right\}, \qquad \text{for any } x_i \in \mathcal{X} \text{ and } n = 1, ..., N, \tag{8.11}$$

which is understood as the proportion of those individuals labelled by $x_i \in \mathcal{X}$ that survive to age $x_i + T_n$. Throughout, we write S_T to denote the derivative of S in its second variable, which is given by

$$S_T(x_i, T_n) = -S(x_i, T_n) \mu(T_n, x_i + T_n).$$

The following definition introduces the notation for the term insurance policies that we shall consider later on.

Definition 8.1. For any $x_i \in \mathcal{I}$ and $y_j \in \mathcal{J}$, the discounted risk-free value at time $t \in I$ of a term insurance policy owned by an individual labelled by $x_i \in \mathcal{X}$ and with term $y_j \in \mathcal{Y}$ is given by

$$z_{ij}(t) = -\sum_{n=1}^{N} S_T(x_i, T_n) \tilde{p}_n(t) \mathbb{1}_{\{T_n \leq y_j\}}. \tag{8.12}$$

Any linear combination of the processes (8.12) gives the risk-free discounted value of a term insurance portfolio. This is stated by the following definition.

Definition 8.2. We call *term insurance portfolio* relative to $\mathcal{X}$ and $\mathcal{Y}$ any matrix $v = \{v_{ij} : \; x_i \in \mathcal{X} \text{ and } y_j \in \mathcal{Y}\}$. Moreover, the discounted risk-free value $v(t)$ of v at time $t \in I$ is defined as

$$v(t) = \sum_{ij} z_{ij}(t) v_{ij}. \tag{8.13}$$

Given any policy portfolio v relative to $\mathcal{X}$ and $\mathcal{Y}$, we regard any of its components v_{ij} as the amount of policies in v that are owned by the class of individuals labelled by $x_i \in \mathcal{X}$ and with term $y_j \in \mathcal{Y}$.

We define the dimension of a term insurance portfolio as the amount of policies that are stored in it.

Definition 8.3. We call *dimension* of a term insurance portfolio v relative to $\mathcal{X}$ and $\mathcal{Y}$ the quantity

$$\dim(v) \triangleq \sum_{ij} v_{ij}.$$

In what follows, for any couple of term insurance policy portfolios v_1 and v_2 relative to $\mathcal{X}$ and $\mathcal{Y}$, we shall write $(v_1 - v_2)(t) \triangleq v_1(t) - v_2(t)$, for any $t \in I$.

Model points risk functional representation. Here and in the sequel, we shall always assume a policy portfolio v relative to $\mathcal{X}$ and $\mathcal{Y}$ to be given and we fix a set $\mathcal{W}$ of term insurance policy portfolios relative to $\mathcal{X}$ and $\mathcal{Y}$. We refer to any $w \in \mathcal{W}$ as a model points portfolio.

According to this setup, Definition 7.2 reduces as follows.

Definition 8.4. We refer to the functional $\mathcal{V} : \mathcal{W} \to \mathbb{R}$ defined by

$$\mathcal{V}(w) = \int_{\mathcal{I}} \mathbb{E}|(v-w)(t) - \mathbb{E}(v-w)(t)|^2 dt, \qquad \text{for any } w \in \mathcal{W},$$

as *the model points risk functional* induced by v over $\mathcal{W}$.

If not otherwise specified, where a term insurance portfolio v is given, we always write $\mathcal{V}$ to denote the model points risk functional induced by v over $\mathcal{W}$.

According to the arguments of the previous chapters, we shall understand the quantity $\mathcal{V}(w)$ as the error that occurs when the policy portfolio v is substituted by the model points portfolio $w \in \mathcal{W}$. In this respect, notice that in the special case when $v \in \mathcal{W}$, one has that $\mathcal{V}(v) = 0$.

According to Definition 6.5, a model points portfolio $w^* \in \mathcal{W}$ is said to be $\mathcal{V}$-optimal relative to v when the following inequality holds true

$$\mathcal{V}(w^*) \leq \mathcal{V}(w), \qquad \text{for any } w \in \mathcal{W}. \tag{8.14}$$

We regard any model points portfolio $w^* \in \mathcal{W}$ satisfying the inequality (8.14) as an optimal representation of the policy portfolio v according to the functional $\mathcal{V}$.

The following result provides an alternative representation of the model points risk functional induced by v over $\mathcal{W}$.

Proposition 8.1. *The model points risk functional induced by a portfolio v over $\mathcal{W}$ admits the following form*

$$\mathcal{V}(w) = \mathbb{E}\bigg\{ \int_I \bigg\| \sum_n \bigg\{ \sum_{ij} (v_{ij} - w_{ij}) S_T(x_i, T_n) \mathbb{1}_{\{T_n \leq y_j\}} \bigg\} \varepsilon_n(t) \tilde{p}_n(t) \bigg\|_H^2 (1-t) dt \bigg\}, \qquad \textit{for any } w \in \mathcal{W}. \tag{8.15}$$

The proof of Proposition 8.1 is based upon similar arguments as those ones provided in the proof of Proposition 7.2.

Proof. Let $E = \mathbb{R}^N$ and for any $t \in I$, set $\tilde{p}(t) \triangleq (\tilde{p}_1(t), ..., \tilde{p}_N(t))$. Thus, let $U = (\mathcal{X} \times \mathcal{Y})^{\mathbb{R}}$ be the class of real matrices $r = \{r(x_i, y_j) : i \in \mathcal{I}, \text{ and } j \in \mathcal{J}\}$.

Consider the functional $Z \in \mathcal{L}(E, U)$ that for any $x_i \in \mathcal{X}$ and $y_j \in \mathcal{Y}$ is given as follows,

$$Z(q)(x_i, y_j) \triangleq -\sum_n q_n S_T(x_i, T_n) \mathbb{1}_{\{T_n \leq y_j\}}, \qquad \text{for any } q \in E,$$

and note that the process $z(t) = \{z(t) : t \in I\}$ defined by the identity (8.13) is recovered when setting

$$z(t) = Z(\tilde{p}(t)), \qquad \text{for any } t \in I.$$

Since Z is linear, for any $q \in E$, its Frechét derivative $\nabla Z(q) \in \mathcal{L}(E, U)$ satisfies the following identity, for any $x_i \in \mathcal{X}$ and $y_j \in \mathcal{Y}$,

$$(\nabla Z(q)q')(x_i, x_j) = -\sum_n q'_n S_T(x_i, T_n)\mathbb{1}_{\{T_n \le y_j\}}, \qquad \text{for any } q' \in E.$$

On the other hand, for any $t \in I$, when letting $\varepsilon(t)$ be the matrix with the nth row given by $\tilde{p}_n(t)\varepsilon_n(t)$, where $\varepsilon_n(t)$ is defined in (8.6), we have that for any $x_i \in \mathcal{X}$ and $y_j \in \mathcal{Y}$, the following identity holds

$$(\nabla Z(\cdot)\varepsilon(t))(x_i, x_j) = -\sum_n \tilde{p}_n(t)\varepsilon_n(t)S_T(x_i, T_n)\mathbb{1}_{\{T_n \le y_j\}}. \tag{8.16}$$

Consider now the function $\zeta : I \times E \to U$ defined as

$$\zeta(t, q) \triangleq Z(q), \qquad \text{for any } t \in I \text{ and } q \in E. \tag{8.17}$$

In this respect, it is worth to be highlighted that the identification (8.17) is allowed since the survival index (8.11) does not depend on $t \in I$, which is the case when imposing the condition (8.10).

Note that the function ζ is of class $\mathscr{C}_b^{1,2}$ and it turns out to be a BL-function relative to $\tilde{p}$. Since (8.5) may be restated as the following E-valued dynamics

$$d\tilde{p}(t) = -\varepsilon(t)dW(t), \tag{8.18}$$

then Proposition 7.1 applies and gives

$$\mathcal{V}(w) = \mathbb{E}\left\{ \int_I \left\| \sum_{ij}(v_{ij} - w_{ij})(\nabla Z(\tilde{p}(t))\epsilon(t))(x_i, x_j) \right\|_H^2 (1-t)dt \right\}, \qquad \text{for any } w \in \mathcal{W}. \tag{8.19}$$

which, jointly with the identity (8.16), gives the representation (8.15). □

8.3 NUMERICAL STUDIES

In this section, we fix a term insurance portfolio v and a set $\mathcal{W}$ of model points. Then, we discuss the numerical treatment of the global optimization problem

$$w^* = \operatorname{argmin}_{w \in \mathcal{W}} \mathcal{V}(w), \tag{8.20}$$

where the functional $\mathcal{V}$ is given by the representation (8.15).

Optimization. Deterministic optimization algorithms such as gradient local optimization schemes are generally computationally efficient. Examples of this kind of algorithms are Pattern Search or the gradient based algorithms like NCG, BFGS, L-BFGS [123] or L-BFGS-B [35]. On the other hand, when considering global optimization problems, local optimization schemes might converge to local minima of the problem.

This problem may be addressed by considering a stochastic optimization algorithm, such as Simulated Annealing (SA). Further details of this matter may be found in [77] and the references therein. Notwithstanding the convergence of the stochastic optimization algorithms is generally slow, they have the advantage to discard local optimal points and avoid to stuck in local solutions.

One technique to obtain faster global optimization algorithms is to consider hybrid algorithms, which combine stochastic optimization algorithms with gradient local optimization schemes. One example of these kind of algorithms is the Basin Hopping algorithm (BH), in which SA is used for sampling the space by randomly generating neighbours, and local gradient algorithms are used to capture the minima starting from the generated points of the stochastic sampler. Details of thes issues may be found in [78, 181] and the references therein.

Parallelization. The functional $\mathcal{V}(w)$ given by (8.15), for any model points portfolio $w \in \mathcal{W}$, has been estimated as a combination of the Monte Carlo simulation for computing the expectation in the measure $\mathbb{P}$, jointly with the discretization of the integral related to the time variable $t \in I$. In particular, a parallel algorithm for multi-CPU architecture and implemented in C++ by considering the OpenMP API has been performed for parallelizing Monte Carlo and speed up the computation.

Table 8.1 shows the performance of the multi-CPU implementation for the calculation of the model points risk functional. The details of the simulation are collected in the caption of the figure.

N. cores	Time (seconds)	Speedup
1	166.64	-
2	97.45	1.71
4	43.62	3.82
8	21.15	7.88
16	10.56	15.78

Table 8.1: Time in seconds and speedups when consideirng 10^4 paths in the Monte Carlo simulation. Here, the term insurance portfolio v has dimension $\dim(v) = 10^3$ and each $w \in \mathcal{W}$ satisfies $w_{ij} = 0$, for any $(x_i, y_j) \notin \mathcal{Z}$, where $\mathcal{Z} \subset \mathcal{X} \times \mathcal{Y}$ is such that its cardinality $|\mathcal{Z}| = 45$.

Details of the simulation. For any fixed $n = 1, ..., N$, the path associated to the dynamics of the forward rate F_n relative to the LIBOR Market Model (8.3) has been simulated by considering the following scheme,

$$\begin{aligned}\ln \hat{F}_n(t+\Delta t) = \ln \hat{F}_n(t) + \sigma_n(t) \sum_{m=1}^{n} \frac{\varrho_{nm}\delta_m\sigma_m(t)\hat{F}_m(t)}{1+\tau_m\hat{F}_m(t)}\Delta t \\ - \frac{\sigma_n(t)^2}{2}\Delta t + \sigma_n(t)\Delta\hat{W}_n(t),\end{aligned} \tag{8.21}$$

where, for any $t \in I$, we write $\hat{F}_n(t)$ and $\Delta\hat{W}_n(t)$ to denote the approximation to $F_n(t)$ and $dW_n(t)$ respectively.

Throughout, any numerical example has been performed by considering the tenor structure $\mathscr{T} \triangleq \{1, 2, ..., 100\}$. For sake of simplicity, we set $F_n(0) = 0.05$, for any $n = 5, ..., 100$ and $F_1(0) = 0.01$, $F_2(0) = 0.02$, $F_3(0) = 0.03$ and $F_4(0) = 0.04$. Moreover, for each $n = 1, ..., N$, we set $\sigma_n(t) = \sigma$, for any $t \in I$, where $\sigma = 0.1$.

On the other hand, the functions $a(s)$ and $b(s)$, for $s \geq 1$, which give the force of mortality (8.9) are defined in such a way that $a(s) = a$ and $b(s) = b$ for any $s \geq 1$, where $a = 0.0003$ and $b = 0.06$.

Tests. Next, we discuss two tests that have been considered in order to study and validate the numerical solution problem (8.20). All the cited tables and figures are commented and reported at the end of this section.

Test 1. In this example, we assume $\mathcal{X}$ and $\mathcal{Y}$ to have the same cardinality $|\mathcal{X}| = |\mathcal{Y}| = 10$. Fix a term insurance portfolio v' relative to $\mathcal{X}$ and $\mathcal{Y}$ and let $\mathcal{W}$ be the class of square real matrices w of order 10, such that $w_{ij} = 0$ in the case when $v'_{ij} = 0$. Thus, given a positive constant $\alpha > 1$, consider the term insurance portfolio v defined as $v \triangleq \alpha v'$. As a result, letting $\mathcal{V}$ be the model points risk functional induced by v over $\mathcal{W}$, since $v \in \mathcal{W}$ and $\mathcal{V}(v) = 0$, one has that the optimal model points portfolio w^* given by the minimization problem (8.20) is expected to approximate v.

Table 8.2 reports the portfolio v' compared with the obtained solution w^* of the problem (8.20), when using the BH algorithm with L-BFGS-B as the local optimizer. The optimization time was 900 seconds (about 15 minutes) and the minimum is given by $\mathcal{V}(w^*) = 2.55357 \times 10^{-08}$. The convergence of the method is shown in Figure 8.1.

On the other hand, the convergence of the optimization procedure using the SA algorithm is shown in Figure 8.2. The time needed was 36000 seconds (about 10 hours).

For this test, Monte Carlo simulations have been performed by considering 10^4 paths.

Test 2. In this test, we fix $\mathcal{X} = \{30, 31, 32, ..., 75\}$ and $\mathcal{Y} = \{5, 6, 7, ..., 30\}$ and thus we fix a term insurance portfolio v relative to $\mathcal{X}$ and $\mathcal{Y}$ such that $\dim(v) = 10^4$. Moreover, we set $\mathcal{X}_{\mathcal{W}} = \{30, 35, 40, ..., 75\}$ and $\mathcal{Y}_{\mathcal{W}} = \{5, 10, 15, ..., 30\}$. Thus, let $\mathcal{W}$ be the family of term insurance portfolios relative to $\mathcal{X}$ and $\mathcal{Y}$, such that $w_{ij} = 0$ for any $(x_i, y_j) \notin \mathcal{X}_{\mathcal{W}} \times \mathcal{Y}_{\mathcal{W}}$.

In this case, the solution $w^* \in \mathcal{W}$ obtained by considering the BH method using L-BFGS-B as local optimizer is shown in Table 8.3. Since $|\mathcal{X}_{\mathcal{W}}| = 9$ and $|\mathcal{Y}_{\mathcal{W}}| = 5$, the solution $w^* \in \mathcal{W}$ is represented as a real 9×5 matrix of the form $w^* = \{w^*_{ij} : x_i \in \mathcal{X}_{\mathcal{W}},\ \text{and } y_j \in \mathcal{Y}_{\mathcal{W}}\}$.

The total computing time needed by the BH algorithm was 9 hours and 36 seconds using 16 processors for the parallel evaluation of the functional $\mathcal{V}(w)$, for $w \in \mathcal{W}$. The result of the optimization is $\mathcal{V}(w^*) = 0.0211981$.

The global property of the hybrid algorithm turns out to be of great importance in this optimization procedure, as several iterations were needed to reach the minimum. Also the convergence speed and accuracy of the L-BFGS-B local optimizer is a key point, as SA was not able to converge in this case. The convergence of the algorithm in this test is shown in Figure 8.3.

$\mathcal{X}$	$\mathcal{Y}$	v'
20.0	50.0	50000
25.0	45.0	100000.0
30.0	40.0	150000.0
35.0	35.0	200000.0
40.0	30.0	250000.0
45.0	25.0	300000.0
50.0	20.0	350000.0
55.0	15.0	400000.0
60.0	10.0	450000.0
65.0	5.0	500000.0

$\mathcal{X}$	$\mathcal{Y}$	w^*
20.0	50.0	50002000
25.0	45.0	99996900
30.0	40.0	150000000
35.0	35.0	200003000
40.0	30.0	249997000
45.0	25.0	300001000
50.0	20.0	350000000
55.0	15.0	400000000
60.0	10.0	450000000
65.0	5.0	500000000

Table 8.2: Comparison between the insurance portfolio $v' \in \mathcal{W}$ on the left and the solution w^* of the minimization problem (8.20) on the right, for $\alpha = 10^3$ for test 1. The optimization has been performed by using the BH algorithm with L-BFGS-B.

	5	**10**	**15**	**20**	**25**
30	45378.4	115912	125831	234832	234832
35	308348	346048	401661	449384	500195
40	1.45884e+09	8.55499e+08	9.26091e+08	7.49281e+08	7.84661e+08
45	6.23481e+07	1.21281e+09	1.33923e+09	4.87623e+08	1.0548e+09
50	1.01028e+09	9.76963e+08	8.46413e+07	1.03506e+09	5.81495e+08
55	7.57129e+08	6.60263e+08	7.68891e+08	3.98609e+08	1.03456e+09
60	1.53556e+09	2.16747e+08	3.15092e+08	5.20125e+08	6.59192e+08
65	1.20964e+09	1.47117e+09	1.47283e+09	5.20125e+08	6.59192e+08
70	9.67702e+08	2.22518e+08	1.38219e+09	1.32217e+09	6.91588e+08

Table 8.3: Solution $w^* \in \mathcal{W}$ of the optimization problem (8.20), by considering the Basin-Hopping algorithm, for test 2.

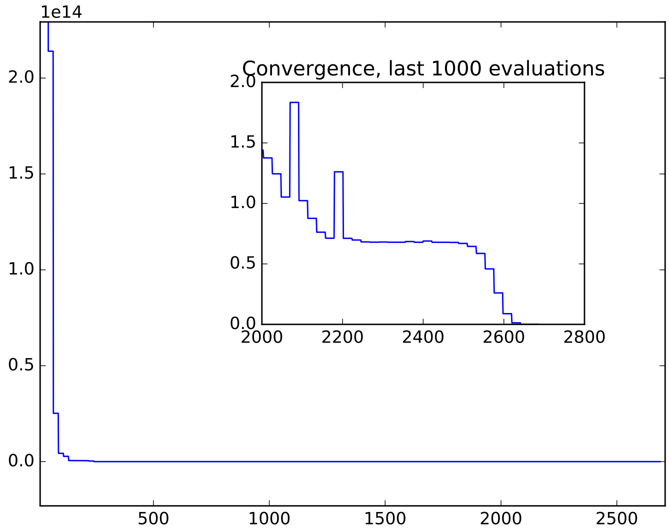

Figure 8.1: Convergence of the BH method with L-BFGS-B as the local optimizer in test 1.

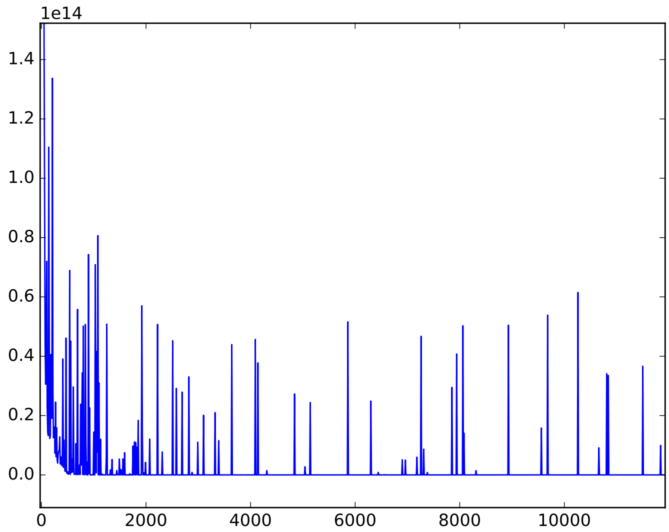

Figure 8.2: Convergence of the SA algorithm in test 1.

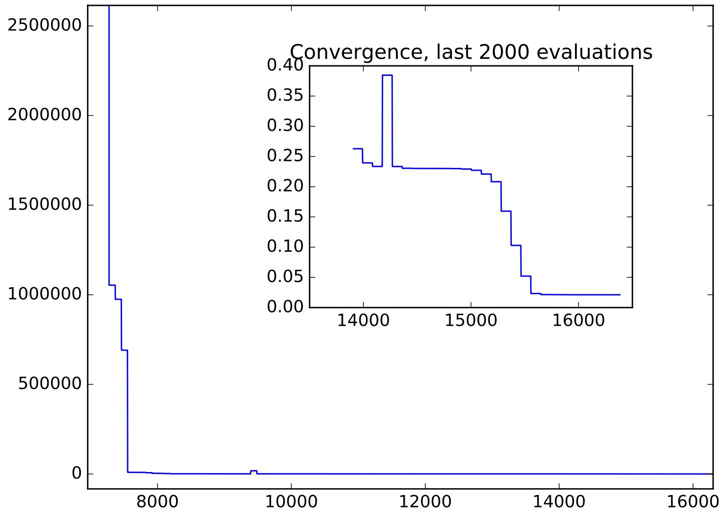

Figure 8.3: Convergence of the BH algorithm with L-BFGS-B as the local optimizer in test 2.

Open Issues and Further Research Lines

The following issues are out of the scope of this dissertation and they represent interesting questions that could be considered for further developments of this work.

MODEL RISK

Serious deficiencies in risk model specification and changes in pricing methodologies were discovered to play a central role within the 2008 financial crisis. One can find details of this matter in [99, 105] and [138].

Model risk or *model uncertainty* refer to the potential losses that a financial institution may suffer due to the development and the application of inaccurate product valuations or an incorrect calibration of the parameters involved. Recent updates in the framework of banking regulation and supervision require financial institutions to explicitly assess the impact of model risk to the product valuations and to deal with it separately from other source of risks. Details of this matter may be found in the documents released by the Bank for International Settlement [17, 18], the European Banking Authority [66] and the Federal Reserve [154].

In recent years, a number of novel approaches have been suggested to assess model uncertainy. Cont [43] developed a framework in which model uncertainty for pricing models is represented as a mapping that satisfies minimal desirable properties, which are derived from the axioms for convex and coherent measures of risk. In the work by Kerkhof et al. [109] model risk is integrated into the computation of the capital reserve and it is assessed as the worst case value at risk within a given class of models. Glasserman and Xu [91] and Breuer and Csiszár [24] address model risk as the deviation from a baseline model by considering relative entropy as models distance. Absolute, relative and local measures for model risk are considered by [9] for the capital calculation. Detering and Packham [59] treat model uncertainty when dealing in the context of hedging contingent claims. In particular, they show that a functional satisfying the assumptions for model uncertainty considered by Cont [43] may be obtained by minimizing the squared deviation related to the overall losses of the hedged position, for different dynamic hedging strategies varying within a certain acceptance set.

In this particular context, the risk functional that we have introduced in order to address the problem of portfolio representation may be understood as the time average of a measure of model uncertainty within a given time horizon. In this regard, it would be desirable to investigate this connection on a deepest level. As a result, the alternative formulations that we provided for the risk functional in portfolio representation may be considered to address model uncertainty in a profitable way for applications in risk management.

MORTALITY RISK

The interest rate term structure and the mortality trend of the population are two of the main factors when dealing with the valuation and the risk management of life insurance products. In this dissertation, the problem of the optimal model points selection for life insurance portfolios has been addressed by considering the changes in interest rates over time as the main source risk. This means that no stochastic fluctuation affecting the time evolution of the mortality rate has been taken into account, and a deterministic survival trend over time for any cohorts is considered.

On the other hand, evidences that the trend of the life expectancy displays a stochastic behaviour have begun to be recognized over the past decades. Details of this matter may be found in the work by Cairns et al [36] and in the references therein. One may refer either to *mortality* or *longevity risk* to denote the possibility faced by the insurer to suffer losses due to unexpected changes in the long-terms life trend displayed by the policy owners.

Based on several similarities between the force of mortality and interest rates [15, 48, 135], many authors proposed a number of stochastic models to describe the evolution of the mortality rate over time. We mainly refer to the Lee-Carter model [122] and its further developments [26, 153], the so-called P-splines models [46, 47] and the model proposed by Cairns et al. [37]. We also refer to the work by Cairns et al. [38], for a qualitative and quantitative comparison of different stochastic mortality models, by discussing their ability to capture historical trends of mortality.

Stochastic mortality models are also crucial when pricing mortality derivatives and survival bonds. This class of instruments have been proposed to hedge the longevity risk when dealing with portfolios of annuities [11, 19, 58, 133, 164, 188]. We refer to [20] for a detailed discussion of this type of market.

As a direct consequence, a more sophisticated approach for the model points selection should be obtained by considering the stochastic time evolution of both the interest rate term structure and the mortality rate. In this context, the dependence between mortality risk and interest rate risk cover a central issue. The first intent to model this form of dependence may be traced back to the work by Jalen and Mamon [103]. Other more recent developments may be found in [55] and [124].

LAPSE RISK

Typical life insurance contracts may involve a surrender option. This gives the policy owner the right to decrease or suspend the insurance cover in whole or in part during the lifetime of the contract. Details of this matter may be found in [68].

An important factor when pricing insurance products is recovered by the so-called *lapse risk*, which is linked to the possibility that the policy owner decides to terminate the contract earlier then what expected by the insurer. Massive early surrenders may constitute a primary liquidity threat for insurance companies and force them to sell assets. Moreover, since surroundings have been recognized to be positively correlated to the level of the interest rates and stocks returns [50, 117], this assets sale might take place in the most inappropriate moment.

In addition to changes in interest rates, there are many other economic factors such as high unemployment and recession that may trigger massive surroundings. As described by Loisel et al. [125], contagion effects like widespread panic or circulating rumor may increase the lapse rate within a certain community. In the work by Barsotti et al. [10] the lapse intensity follows a dynamic contagion process as introduced in the work by Dassios and Zhao [51] and the interest rate dynamics is considered as a specific factor affecting the lapse decision. For other details about lapse risk we refer to [10, 70, 126, 187] and the references therein.

Analysis of lapse risk and its treatment in model points selection is out of the scope of this thesis. However, they could represent an interesting feature for further development of this work.

Resumen extenso

Esta tesis es el resultado de las actividades de investigación realizadas como estudiante de doctorado en el Departamento de Matemáticas de la Universidade da Coruña en el período comprendido entre octubre de 2015 y septiembre de 2018. Este trabajo ha sido desarrollado bajo los auspicios del proyecto Wakeupcall, financiado en el marco de la UE Horizon2020 dentro del programa ITN-MSCA, que ha sido promovido con el fin de contribuir en los avances que se han producido dentro de las prácticas financieras como consecuencia de la crisis financiera de 2008.

Muchos de estos avances propiciaron el estudio de una serie de problemas en la industria de seguros de vida, que involucran el desarrollo de modelos sofisticados. El mercado actual de seguros de vida es mucho más complejo que en el pasado y ofrece una amplia gama de contratos diferentes, que son difíciles de distinguir correctamente sin un conocimiento experto. Como resultado de estos desarrollos, el ajuste (matching) entre los flujos de efectivo de los activos y pasivos ha adquirido una relevancia significativa en el sector de los seguros, y además se ha vuelto progresivamente más conectado con los mercados financieros. Hoy en día, las compañías de seguros tratan con una gran cantidad de activos diferentes, que se negocian con el fin de cubrir su exposición. Incluso a nivel regulatorio, el sector de seguros de vida experimentó transformaciones drásticas destinadas a promover el bienestar general. Con el fin de supervisar y hacer cumplir la regulación dentro del sector de seguros de la Unión

Europea, la European Insurance and Occupational Pensions Authority (EIOPA) se creó en 2010 como parte del Sistema Europeo de Supervisión Financiera. Constituye una entidad independiente que asesora a la Comisión Europea, al Parlamento Europeo y al Consejo de la Unión Europea. Algunos cambios en la regulación dentro del sector de seguros de vida también ha sido promovido por el Tribunal de Justicia de las Comunidades Europeas, donde, por ejemplo, se resolvió el caso principal de Test Achats en marzo de 2011, que prohíbe la discriminación basada en el género en los precios de las pólizas.

Según la Directiva Solvencia II emitida por la Unión Europea en noviembre de 2009, que entró en vigor en enero de 2016, las compañías de seguros deben almacenar una cantidad suficiente de fondos destinados a reducir el costo social de las posibles pérdidas y a prevenir las posibles futuras crisis. El capital requerido refleja el riesgo que afronta la aseguradora y debe basarse en un enfoque de valoración razonable de los activos y los pasivos en el balance general. Según este marco, la gestión del riesgo de mercado también desempeña un papel central en el sector de seguros. A este respecto, la variación de la estructura temporal de los tipos de interés proporciona una importante fuente de riesgo, ya que generalmente la mayoría de los productos comercializados por los aseguradores para cubrir su exposición abarcan desde bonos o instrumentos de esa naturaleza hasta productos complejos que dependen del tipo interés.

El documento de la tesis se ha dividido en dos partes.

La primera parte trata sobre temas relacionados con los estudios realizados sobre la teoría moderna de las medidas de riesgo. Esta teoría comenzó con el trabajo inicial de Artzner et al. [7], que ha tenido un gran desarrollo posterior en la última década. Según dicho marco, el riesgo de caída o de pérdida de una cierta exposición se evalúa sobre la base de la información pasada. Como consecuencia, al ampliar el conjunto de datos históricos, la estabilidad asintótica de los estimadores de riesgo proporciona una herramienta fundamental. En este contexto, además de la consistencia estadística, en Cont et al. [44] se destacó que la noción de robustez cualitativa introducida por

Hampel [97, 98] y Huber [102] es una propiedad deseable de la estabilidad asintótica cuando se trata de estimadores de riesgo.

Se refiere a la propiedad del estimador de riesgo asociado de ser estable con respecto a pequeños cambios que puedan afectar a la distribución de la exposición financiera. Nos referimos principalmente a los trabajos de Krätschmer, Schied and Zähle [114, 115], en los que se introduce una noción de robustez cualitativa cuando se considera una sucesión de variables aleatorias independientes e igualmente distribuidas. Por otro lado, los trabajos de Zähle [186, 184, 185] también son notables y proporcionan una caracterización de la robustez cuando se trata de sucesiones aleatorias que muestran una estructura de dependencia interna. Además, en los trabajos citados, todos los resultados obtienen al considerar un refinamiento de la medida de la topología débil en términos de una determinada función de evaluación, gauge, que mide el desplazamiento de las distribuciones de sus colas. Este planteamiento es razonable en la estimación del riesgo, ya que la medida de la topología débil carece de sensibilidad con respecto a las colas de las distribuciones. Básicamente, esto significa que dos leyes pueden estar bastante cerca con respecto a alguna métrica en relación con la topología débil y al mismo tiempo mostrar un comportamiento completamente diferente en sus colas. Por otro lado, muchos funcionales de riesgo comunes no son cualitativamente robustos con respecto a la medida de la topología débil.

Uno de los objetivos principales del trabajo ha sido el desarrollo de una versión refinada de la robustez cualitativa que se aplica cuando se trata de sucesiones aleatorias estacionarias, que son aquellas sucesiones cuya ley resulta invariante bajo la acción del operador de desplazamiento (shift). Para lograr este objetivo, uno de los principales problemas era el estudio de la convergencia de los procesos empíricos asociados a una sucesión dada de variables En particular, cuando se trata de sucesiones estacionarias, el límite del proceso empírico se ha caracterizado en términos de una medida de probabilidad aleatoria específica que, además, puede considerarse como una variable aleatoria que toma valores en una familia específica de leyes. La noción propuesta de robustez se ha presentado en términos de la acción de un endomorfismo genérico, que codifica una perturbación específica en el espacio muestral, que se define como la familia completa de sucesiones muestrales. Como resultado directo, obtuvimos que esta formulación puede interpretarse como el hecho de que las pequeñas perturbaciones al nivel del conjunto de datos solo producen pequeñas perturbaciones en términos de la ley asintótica de la familia de estimadores. Por otro lado, una de las aportaciones principales del trabajo es la noción estadística refinada de robustez cualitativa en términos del refinamiento topológico mencionado anteriormente. Además, al tratar con medidas aleatorias, la σ-álgebra de Borel generada por la medida de la topología débil posee propiedades deseables que juegan un papel central. Por lo tanto, la caracterización de la σ-álgebra de Borel generada por la topología refinada citada anteriormente ha representado una característica principal para formular todo el resultado de una manera adecuada.

A nivel de inferencia estadística, el resultado que presentamos admite también una interpretación natural en términos de análisis bayesiano, ya que permanece en vigor cuando consideramos aleatorias intercambiables, que son aquellas sucesiones cuya ley es invariante bajo la acción de permutaciones de orden finito. De hecho, cualquier sucesión estacionaria también es intercambiable. La intercambiabilidad proporciona el pilar principal del enfoque bayesiano para el análisis inferencial. Más concretamente, cuando se trata del marco no paramétrico, la ley inducida por cualquier medida aleatoria puede considerarse como la distribución previa del modelo estadístico de referencia. La distribución a priori es el pilar principal de las estadísticas bayesianas y representa el conocimiento subjetivo antes de que se lleve a cabo el experimento. De acuerdo con dicha formulación, el resultado discutido puede considerarse como una forma de estabilidad obtenida cuando la distribución previa del modelo se ve obligada a cambiar de tal manera que se satisfagan ciertas condiciones débiles. Además, destacamos una conexión interesante entre la noción propuesta de robustez y la de *elicitabilidad*, que proporciona un aspecto ampliamente discutido en la evaluación de predicciones puntuales, ya que puede considerarse como la forma dual de optimalidad en el marco de la predicción puntual. Como consecuencia de ello, muchos autores

[81, 93, 142, 189] creen que la *elicitabilidad* proporciona una herramienta natural para realizar *back-testing*.

El principal objetivo de la segunda parte de la tesis es el estudio del problema de la reducción del tamaño de una cartera de pólizas de seguros con el objetivo de su análisis y gestión, de modo que los riesgos inherentes se mantengan correctamente representados al reducir el tamaño. A las compañías europeas de seguros se les ha requerido que evalúen sus carteras así como que realicen los análisis de sensibilidad para demostrar la validez de sus modelos, considerando proyecciones de flujos de caja en una aproximación póliza a póliza. Además, se les permite calcular estas proyecciones reemplazando cualquier grupo homogéneo de pólizas por un contrato representativo adecuado, lo que se conoce como *model points*. Este procedimiento se utiliza para acelerar el proceso, que se lleva a cabo dirariamente, ya que la complejidad de tratar toda la cartera puede conducir a tiempos de cálculo muy extensos y prohibitivos. Esta técnica de reemplazar la toda la cartera por otra cartera está permitida siempre que la estructura del riesgo subyacente de la cartera inicial se mantenga en la nueva cartera.

En este trabajo hemos comenzado por caracterizar este problema como el de la sustitución eficiente de una cartera por otra simple con riesgos similares. Además, es razonable asumir que la cartera más simple debe verificar ciertas restricciones. Hemos llamado a este problema con el término de representación de carteras. Por ejemplo, encontramos este problema cuando definimos una estrategia de cobertura sujeta a restricciones de política y presupuesto. Otro ejemplo de este problema surge cuando se busca reducir la escala, y por tanto la complejidad, de una cartera específica con el objetivo de analizarla y gestionarla, sin perder la representación de su estructura de riesgos inherentes. Abordamos este problema definiendo un concepto razonable de optimalidad basado en un criterio específico de comparación. En concreto, presentamos una aproximación basada en la minimización de cierto funcional de riesgo, que mide la discrepancia media entre la cartera original y cualquier cartera candidata entre las más simples que representan su exposición. En particular, el riesgo se ha medido en función de los factores de riesgo subyacentes dentro de un horizonte temporal.

En las pasadas décadas, muchos resultados de integración estocástica en espacios de Hilbert se han generalizado al marco de los espacios de Banach. Una teoría de la integración estocástica para procesos tomando valores en espacios de Banach se ha propuesto inicialmente para espacios con martingalas de tipo 2, tal y como se presenta en los trabajos desarrollados en Dettweiler [60, 61], Neidhardt [139] y Ondreját [141]. Algunos de los autores citados han desarrollado la teoría de integración estocástica en los llamados espacios de Banach 2-uniformemente regulares. Además, en Pisier [145] se prueba que la propiedad de martingala de tipo 2 es equivalente a la regularidad 2-uniforme. Brzeźniak [28, 29, 30, 27] continuó desarrollando las integrales estocásticas de Neidhardt y Dettweiler proporcionando algunas aplicaciones a la teoría de las ecuaciones diferenciales parciales estocásticas. En una línea diferente, Garling [88] y McConnell [134] explotaron la definición probabilística de los llamados espacios de Banach UMD introducidos por Burkholder [32] y Bourgain [23] para proponer una teoría de cálculo estocástico en esta clase de espacios. Las ideas propuestas por Garling y McConnell han sido extendidas ampliamente por van Neerven, Veraar y Weis [173, 174, 177], obteniendo algunas aplicaciones muy relevantes de esta teoría al estudio de las ecuaciones en derivadas parciales estocásticas [170, 176, 175, 31].

Motivados por estos desarrollos, algunos autores generalizaron muchos de los conceptos habituales del cálculo de Malliavin al marco de los espacios de Banach. En la presente tesis hemos revisado el trabajo de Maas y van Neerven [128], en el cual la fórmula de representación de Clark-Ocone se presenta cuando tratamos con variables aleatorias que toman valores en un espacio de Banach de tipo UMD.

En línea con los avances descritos anteriormente, en esta tesis hemos modelado las dinámicas de los factores de riesgo como procesos de difusión en un cierto espacio de Banach y después proponemos dos formulaciones diferentes del funcional antes citado en el marco de la teoría de integración estocástica en espacios UMD desarrollada por Van Neerven, Veraar y Weis [173, 174, 177].

La primera formulación se plantea en términos del operador derivada de Malliavin.

En particular, el criterio que hemos obtenido resulta ser similar al planteamiento de minimización sugerido en [110] para abordar el problema de cobertura de carteras de bonos, en los que se maneja un concepto refinado de duración de un bono introducido mediante el cálculo de Malliavin para campos aleatorios gaussianos en el marco de espacios de Hilbert. La duración es una herramienta que se ha utilizado durante mucho tiempo en el ámbito de los mercados de renta fija, ya que proporciona la sensibilidad de una cartera a las fluctuaciones de los tipos de interés.

La segunda formulación se ha obtenido bajo condiciones adicionales sobre el modelo y fundamentalmente involucra la componente difusiva de la dinámica de la curva de descuento. En particular, esta representación admite una interpretación muy interesante en términos de una condición generalizada de consistencia del mercado. En relación con estos desarrollos, la formulación de una teoría de carteras en el marco de un espacio de Banach ha representado un primer objetivo.

En relación con estos desarrollos, un primer objetivo era la formulación de una teoría de carteras en el marco de un espacio de Banach. Un ejemplo interesante en este marco se ha obtenido discutiendo cómo estos argumentos pueden ser aplicados en el mercado de renta fija, cuando se analiza el problema de encontrar una representación óptima de una cartera fija de activos considerando una cartera de bonos cupón cero. En este sentido, consideramos una dinámica infinito dimensional que gobierna la evolución estocástica de la curva de precios. Es más, de acuerdo con los modelos clásicos, las dinámicas infinito dimensionales se consideran para capturar la tendencia estocástica de la estructura de tipos de interés a lo largo del tiempo. Es importante mencionar que muchos autores han desarrollado una teoría de carteras de bonos en el marco de espacios de Hilbert de dimensión infinita (ver [39, 40, 69, 156], por ejemplo).

Como ejemplo particular de esta aplicación, hemos estudiado el caso en el que una cartera de productos dependientes del tipo de interés se representa mediante una cartera compuesta por un único bono cupón cero. Como consecuencia directa, se recupera una noción más refinada de duración cuando se minimiza el funcional de riesgo explicado anteriormente. Este último planteamiento ha resultado muy relevante por su conexión con el problema de selección de *model points* De hecho, cuando se trata con carteras de pólizas, una de las principales fuentes de riesgo surge de las fluctuaciones de los tipos de interés con el tiempo. Nosotros hemos presentado una teoría general de carteras de pólizas en el marco de un espacio de Banach, en el cual la cuestión de la selección óptima de *model points* se ha evaluado en relación al problema de representación de carteras previamente analizado desde un punto de vista teórico.

Se ha puesto un énfasis especial en el caso de toda la cartera de pólizas de seguros, que por definición pagan una suma de todos los beneficios en caso de fallecimiento del poseedor de la póliza. En este caso particular, el criterio de minimización que hemos obtenido es similar al que se basaba en la duración de la cartera en el mercado de renta fija.

Por otro lado, desde el punto de vista numérico, la minimización del funcional de riesgo que hemos introducido para resolver el problema de representación de carteras puede resultar muy intensivo computacionalmente, ya que normalmente conduce a un problema de minimización global en alta dimensión. Los algoritmos de optimización deterministas, como el método de optimización local del gradiente, resultan ser computacionalmente eficientes. Algunos ejemplos de este tipo de algoritmos son *pattern search* o algoritmos basados en el gradiente como el gradiente conjugado no lineal (NCG), BFGS, L-BFGS (descritos en [123], por ejemplo) o L-BFGS-B (ver [35], por ejemplo). Por otro lado, cuando consideramos problemas de optimización global, los algoritmos de optimización local anteriores pueden terminar convergiendo hacia mínimos locales, que no proporcionan la solución del problema planteado. Este inconveniente puede ser solventado mediante la utilización de algoritmos de optimización estocásticos,como por ejemplo *Simulated Annealing* (SA). Aunque la convergencia de estos algoritmos de optimización estocásticos es lenta en general, tienen la ventaja de que descartan los mínimos locales y evitan el quedarse atrapados en soluciones locales del problema.

Una posibilidad para obtener algoritmos de optimización global más rápidos que ha sido considerada en esta tesis es el uso de algoritmos híbridos, que combinan los algoritmos de optimización estocásticos con los métodos locales de optimización. Un ejemplo de esta clase de algoritmos es el algoritmo de Basin Hopping (BH), en el cual el algoritmo global estocástico de SA se emplea para muestrear el espacio, generando puntos vecinos de manera aleatoria, y se combina con algoritmos locales de tipo gradiente para capturar capturar el mínimo partiendo inicialmente de uno de esos vecinos generados por el muestreo estocástico. Estos algoritmos se analizan en los trabajos [78, 181]. Por otro lado, el coste computacional de los algoritmos resultantes se ha gestionado mediante una paralelización eficiente en las arquitecturas de ordenadores adecuadas.

En relación con estos últimos aspectos computacionales, se ha presentado un estudio numérico para analizar el comportamiento de las metodologías propuestas cuando se aplican a una cartera grande de contratos de seguros asociados a tipos de interés, es decir, los contratos pagan una suma de beneficios en caso de fallecimiento del poseedor de la póliza, siempre que esto ocurra antes de un tiempo determinado. En esta aplicación, se usa un modelo de mercado (LIBOR Market Model, LMM) para simular la evolución estocástica de los tipos forward en el tiempo. El LMM se utiliza mucho en la práctica, pues presenta la ventaja de poder ser calibrado a mercado de modo consistente.

El contenido de esta tesis doctoral se ha organizado en dos partes.

La Parte I contiene las aportaciones que se han llevado a cabo en el marco de la teoría moderna de medidas de riesgo y de estimación robusta del riesgo.

En el Capitulo 2 se revisan brevemente los resultados de convergencia que motivan los problemas que se abordan en los siguientes capítulos. Se pone especial énfasis en la noción de topología débil y en las formas comunes de simetría en probabilidad, como por ejemplo la estacionaridad (stationarity) y la intercambiabilidad (exchangeability).

En el Capitulo 3 se presenta una revisión particular de la teoría moderna de medidas de riesgo, basada fundamentalmente en las referencias que se citan en este capítulo. Se ha dedicado una atención especial al concepto de robustez cualitativa en estimación del riesgo.

El Capítulo 4 se basa fundamentalmente en los contenidos del trabajo [75], excepto la Sección 4.1, que no forma parte de ese trabajo y que representa el desarrollo de diferentes resultados de la literatura, que se encontraron adecuados por razones de completitud del capítulo. En este capítulo se introduce una noción refinada de robustez cualitativa que se deriva a partir del concepto de consistencia estadística en términos de medidas de probabilidad aleatoria.

La Parte II de la tesis trata del problema de representación de carteras, con atención especial a las que aparecen en el sector de seguros de vida.

En el Capítulo 5 se recogen los resultados fundamentales de la teoría de integración estocástica en espacios de Banach, con una cierta revisión sobre la literatura en el tema. Esta revisión se basa en las referencias citadas en este capítulo. Aunque la mayor parte de los resultados son conocidos, en la Sección 5.4 se han analizado y demostrado aquellos que no se encontraron en la forma adecuada en la literatura.

Los Capítulos 6 y 7 incluyen resultados originales basados en el trabajo [79]. En concreto, en el Capítulo 6 se aborda el problema de representación de carteras en el marco de la teoría de integración estocástica en espacios de Banach, mientras que en el Capítulo 7 se analiza este planteamiento cuando se tratan carteras de seguros de vida, haciendo especial énfasis en carteras compuestas de pólizas de seguros de vida (whole life insurance policies).

El Capítulo 8 se basa en el trabajo [76]. En concreto, contiene un estudio numérico de la aproximación planteada en el capítulo anterior cuando se aplica a carteras de seguros a plazo, considerando el LMM como modelo para describir la evolución en el tiempo de la estructura de tipos de interés forward.

Finalmente, se incluye un apartado en el que se analizan algunos problemas abiertos y se plantean ideas de posibles líneas futuras de investigación.

Resumo extenso

Esta tese é o resultado das actividades de investigación realizadas como estudante de doutoramento no Departamento de Matemáticas da Universidade dá Coruña no período comprendido entre outubro de 2015 e setembro de 2018. Este traballo foi desenvolto baixo os auspicios do proxecto Wakeupcall, financiado no marco da UE Horizon2020 dentro do programa ITN- MSCA, que foi promovido co fin de contribuír nos avances que se produciron dentro das prácticas financeiras como consecuencia da crise financeira de 2008.

Moitos destes avances propiciaron o estudo dunha serie de problemas na industria de seguros de vida, que involucran o desenvolvemento de modelos máis sofisticados. O mercado actual de seguros de vida é moito máis complexo que no pasado e ofrece unha ampla gama de contratos diferentes, que son difíciles de distinguir correctamente sen un coñecemento experto. Como resultado destes desenvolvementos, o axuste (matching) entre os fluxos de efectivo dos activos e pasivos adquiriu unha relevancia significativa no sector dos seguros, e ademais volveuse progresivamente máis conectado cos mercados financeiros. Hoxe en día, as compañías de seguros tratan cunha gran cantidade de activos diferentes, que se negocian co fin de cubrir a súa exposición. Mesmo a nivel regulatorio, o sector de seguros de vida experimentou transformacións drásticas destinadas a promover o benestar xeral. Co fin de supervisar e facer cumprir a regulación dentro do sector de seguros da Unión Europea, a

European Insurance and Occupational Pensions Authority (EIOPA) creouse en 2010 como parte do Sistema Europeo de Supervisión Financeira. Constitúe unha entidade independente que asesora á Comisión Europea, ao Parlamento Europeo e ao Consello da Unión Europea. Algúns cambios na regulación dentro do sector de seguros de vida tamén foi promovido polo Tribunal de Xustiza das Comunidades Europeas onde, por exemplo, resolveuse o caso principal de Test Achats en marzo de 2011, que prohibe a discriminación baseada no xénero nos prezos das pólizas.

Segundo a Directiva Solvencia II emitida pola Unión Europea en novembro de 2009, que entrou en vigor en xaneiro de 2016, as compañías de seguros deben almacenar unha cantidade suficiente de fondos destinados a reducir o custo social das posibles perdas e a previr as posibles futuras crises. O capital requirido reflicte o risco que afronta a aseguradora e debe basearse nun enfoque de valoración razoable dos activos e os pasivos no balance xeral. Segundo este marco, a xestión do risco de mercado tamén desempeña un papel central no sector de seguros. A este respecto, a variación da estrutura temporal dos tipos de xuro proporciona unha importante fonte de risco, xa que xeralmente a maioría dos produtos comercializados polas aseguradoras para cubrir a súa exposición abarcan desde bonos ou instrumentos desa natureza ata produtos complexos que dependen do tipo de xuro.

O documento da tese organizóuse en dúas partes.

A primeira parte trata sobre temas relacionados cos estudos realizados sobre a teoría moderna das medidas de risco. Esta teoría comezou co traballo inicial de Artzner et al. [7], que tivo un gran desenvolvemento posterior na última década. Segundo o devandito marco, o risco de perda dunha certa exposición avalíase sobre a base da información pasada. Como consecuencia, ao ampliar o conxunto de datos históricos, a estabilidade asintótica dos estimadores de risco proporciona unha ferramenta fundamental. Neste contexto, ademais da consistencia estatística, en Cont et al. [44] destácase que a noción de robustez cualitativa introducida por Hampel [97, 98] e Huber [102] é unha propiedade desexable da estabilidade asintótica cando se trata de estimadores de risco. Refírese á propiedade do estimador de risco asociado de ser estable con respecto a pequenos cambios que poidan afectar á distribución da exposición financeira. Referímonos principalmente aos traballos de Krätschmer, Schied e Zähle [114, 115], nos que se introduce unha noción de robustez cualitativa cando se considera unha sucesión de variables aleatorias independentes e igualmente distribuídas. Doutra banda, os traballos de Zähle [186, 184, 185] tamén son notables e proporcionan unha caracterización da robustez cando se trata de sucesións aleatorias que mostran unha estrutura de dependencia interna. Ademais, nos traballos citados, todos os resultados obtéñense ao considerar un refinamento da medida da topoloxía feble en termos dunha determinada función de avaliación, gauge, que mide o desprazamento das distribucións nas súas colas. Esta formulación é razoable na estimación do risco, xa que a medida da topoloxía feble carece de sensibilidade con respecto ás colas das distribucións. Basicamente, isto significa que dúas leis poden estar bastante preto con respecto a algunha métrica en relación coa topoloxía feble e ao mesmo tempo mostrar un comportamento completamente diferente nas súas colas. Doutra banda, moitos funcionais de risco comúns non son cualitativamente robustos con respecto á medida da topología feble.

Un dos obxectivos principais do traballo foi o desenvolvemento dunha versión refinada da robustez cualitativa que se aplica cando se trata de sucesións aleatorias estacionarias, que son aquelas sucesións cuxa lei resulta invariante baixo a acción do operador de desprazamento (shift). Para acadar este obxectivo, un dos principais problemas era o estudo da converxencia dos procesos empíricos asociados a unha sucesión dada de variables. En particular, cando se trata de sucesións estacionarias, o límite do proceso empírico caracterizouse en termos dunha medida de probabilidade aleatoria específica que, ademais, pode considerarse como unha variable aleatoria que toma valores nunha familia específica de leis. A noción proposta de robustez presentouse en termos da acción dun endomorfismo xenérico, que codifica unha perturbación específica no espazo muestral, que se define como a familia completa de sucesións muestrales. Como resultado directo, obtivemos que esta formulación pode interpretarse como o feito de que as pequenas perturbacións ao nivel do conxunto de datos só producen pequenas perturbacións en termos da lei asintótica da familia de estimadores. Doutra banda, unha das achegas principais do traballo é a noción estatística refinada de robustez cualitativa, en termos do refinamento topolóxico mencionado anteriormente. Ademais, ao tratar con medidas aleatorias, o σ-álxebra de Borel xerada pola medida da topoloxía feble posúe propiedades desexables que xogan un papel central. Por tanto, a caracterización do σ-álxebra de Borel xerada pola topoloxía refinada citada anteriormente representou unha característica principal para formular todo o resultado dunha maneira axeitada.

A nivel de inferencia estatística, o resultado que presentamos admite tamén unha interpretación natural en termos de análises bayesiano, xa que permanece en vigor cando consideramos sucesións aleatorias intercambiables, que son aquelas sucesións cuxa lei é invariante baixo a acción de permutacións de orde finito. De feito, calquera sucesión estacionaria tamén é intercambiable. A intercambiabilidade proporciona o alicerce principal do enfoque bayesiano para a análise inferencial. Máis concretamente, cando se trata do marco non paramétrico, a lei inducida por calquera medida aleatoria pode considerarse como a distribución previa do modelo estatístico de referencia. A distribución a priori é o alicerce principal da estatística bayesiana e representa o coñecemento subxectivo antes de que se leve a cabo o experimento. Dacordo con dita formulación, o resultado discutido pode considerarse como unha forma de estabilidade obtida cando a distribución previa do modelo vese obrigada a cambiar de tal maneira que se satisfagan certas condicións febles. Ademais, destacamos unha conexión interesante entre a noción proposta de robustez e a de elicitabilidade, que proporciona un aspecto amplamente discutido na avaliación de predicións puntuais, xa que pode considerarse como a forma dual de optimalidade no marco da predición puntual. Como consecuencia diso, moitos autores [81, 93, 142, 189] cren que a elicitabilidade proporciona unha ferramenta natural para realizar back-testing.

O principal obxectivo da segunda parte da tese é o estudo do problema da redución do tamaño dunha carteira de pólizas de seguros co obxectivo da súa análise e xestión, de modo que os riscos inherentes mantéñanse correctamente representados ao reducir o tamaño. Ás compañías europeas de seguros requiríuselles que avalíen as súas carteiras así como que realicen as análises de sensibilidade para demostrar a validez dos seus modelos, considerando proxeccións de fluxos de caixa nunha aproximación póliza a póliza. Ademais, permíteselles calcular estas proxeccións substituíndo calquera grupo homoxéneo de pólizas por un contrato representativo adecuado, o que se coñece como *model points*. Este procedemento utilízase para acelerar o proceso, que se leva a cabo diariamente, xa que a complexidade de tratar toda a carteira pode conducir a tempos de cálculo moi longos e prohibitivos. Esta técnica de substituír toda a carteira por outra carteira está permitida sempre que a estrutura do risco subxacente da carteira inicial mantéñase na nova carteira.

Neste traballo comezamos por caracterizar este problema como o da substitución eficiente dunha carteira por outra simple con riscos similares. Ademais, é razoable asumir que a carteira máis simple debe verificar certas restricións. Chamamos a este problema co termo de representación de carteiras. Por exemplo, atopamos este problema cando definimos unha estratexia de cobertura suxeita a restricións de política e orzamento. Outro exemplo deste problema xorde cando se busca reducir a escala, e por tanto a complexidade, dunha carteira específica co obxectivo de analizala e xestionala, sen perder a representación da súa estrutura de riscos inherentes. Abordamos este problema definindo un concepto razoable de optimalidade baseado nun criterio específico de comparación. En concreto, presentamos unha aproximación baseada na minimización de certo funcional de risco, que mide a discrepancia media entre a carteira orixinal e calquera carteira candidata entre as máis simples que representan a súa exposición. En particular, o risco mediuse en función dos factores de risco subxacentes dentro dun horizonte temporal.

Nas pasadas décadas, moitos resultados de integración estocástica en espazos de Hilbert xeneralizáronse ao marco dos espazos de Banach. Unha teoría da integración estocástica para procesos tomando valores en espazos de Banach propúxose inicialmente para espazos con martingalas de tipo 2, tal e como se presenta nos traballos desenvoltos en Dettweiler [60, 61], Neidhardt [139] e Ondreját [141]. Algúns dos autores citados desenvolveron a teoría de integración estocástica nos chamados espazos de Banach 2- uniformemente regulares. Ademais, en Pisier [145] próbase que a propiedade de martingala de tipo 2 é equivalente á regularidade 2-uniforme. Brzézniak [28, 29, 30, 27] continuou desenvolvendo as integrais estocásticas de Neidhardt e Dettweiler proporcionando algunhas aplicacións á teoría das ecuacións diferenciais parciais estocásticas. Nunha liña diferente, Garling [88] e McConnell [134] explotaron a definición probabilística dos chamados espazos de Banach UMD introducidos por Burkholder [32] e Bourgain [23] para propoñer unha teoría de cálculo estocástico nesta clase de espazos. As ideas propostas por Garling e McConnell foron estendidas amplamente por van Neerven, Veraar e Weis [173, 174, 177], obtendo algunhas aplicacións moi relevantes desta teoría ao estudo das ecuacións en derivadas parciais estocásticas [170, 176, 175, 31]. Motivados por estes desenvolvementos, algúns autores xeneralizaron moitos dos conceptos habituais do cálculo de Malliavin ao marco dos espazos de Banach. Na presente tese tamén revisamos o traballo de Maas e van Neerven [128], no cal a fórmula de representación de Clark- Ocone preséntase cando tratamos con variables aleatorias que toman valores nun espazo de Banach de tipo UMD.

En liña cos avances descritos anteriormente, nesta tese modelamos as dinámicas dos factores de risco como procesos de difusión nun certo espazo de Banach e despois propoñemos dúas formulacións diferentes do funcional antes citado no marco da teoría de integración estocástica en espazos UMD desenvolta por van Neerven, Veraar e Weis [173, 174, 177].

A primeira formulación exponse en termos do operador derivada de Malliavin. En particular, o criterio que obtivemos resulta ser similar á formulación de minimización suxerido en [110] para abordar o problema de cobertura de carteiras de bonos, nos que se manexa un concepto refinado de duración dun bono introducido mediante o cálculo de Malliavin para campos aleatorios gaussianos no marco de espazos de Hilbert. A duración é unha ferramenta que se utilizou durante moito tempo no ámbito dos mercados de renda fixa, xa que proporciona a sensibilidade dunha carteira ás fluctuaciones dos tipos de xuro.

A segunda formulación obtívose baixo condicións adicionais sobre o modelo e fundamentalmente involucra a compoñente difusiva da dinámica da curva de desconto. En particular, esta representación admite unha interpretación moi interesante en termos dunha condición xeneralizada de consistencia do mercado. En relación con estes desenvolvementos, a formulación dunha teoría de carteiras no marco dun espazo de Banach representou un primeiro obxectivo.

En relación con estes desenvolvementos, un primeiro obxectivo era a formulación dunha teoría de carteiras no marco dun espazo de Banach. Un exemplo interesante neste marco obtívose discutindo como estes argumentos poden ser aplicados no mercado de renda fixa, cando se analiza o problema de atopar unha representación óptima dunha carteira fixa de activos considerando unha carteira de bonos cupón cero. Neste sentido, consideramos unha dinámica infinito dimensional que goberna a evolución estocástica da curva de prezos. É máis, de acordo cos modelos clásicos, as dinámicas infinito dimensionais considéranse para capturar a tendencia estocástica da estrutura de tipos de xuro ao longo do tempo. É importante mencionar que moitos autores desenvolveron unha teoría de carteiras de bonos no marco de espazos de Hilbert de dimensión infinita (ver [39, 40, 69, 156], por exemplo).

Como exemplo particular desta aplicación, estudamos o caso no que unha carteira de produtos dependentes do tipo de xuro represéntase mediante unha carteira composta por un único bono cupón cero. Como consecuencia directa, recupérase unha noción máis refinada de duración cando se minimiza o funcional de risco explicado anteriormente.

Esta última formulación resultou moi relevante pola súa conexión co problema de selección de *model points*. De feito, cando se trata con carteiras de pólizas, unha das principais fontes de risco xorde das variacions dos tipos de xuro co tempo. Nós presentamos unha teoría xeral de carteiras de pólizas no marco dun espazo de Banach, no cal a cuestión da selección óptima de *model points* avaliouse en relación ao problema de representación de carteiras previamente analizado desde un punto de vista teórico.

Púxose unha énfase especial no caso de toda a carteira de pólizas de seguros, que por definición pagan unha suma de todos os beneficios en caso de falecemento do posuidor da póliza. Neste caso particular, o criterio de minimización que obtivemos é similar ao que se baseaba na duración da carteira no mercado de renda fixa.

Doutra banda, desde o punto de vista numérico, a minimización do funcional de risco que introducimos para resolver o problema de representación de carteiras pode resultar moi intensivo computacionalmente, xa que normalmente conduce a un problema de minimización global en alta dimensión. Os algoritmos de optimización deterministas, como o método de optimización local do gradiente, resultan ser computacionalmente eficientes. Algúns exemplos deste tipo de algoritmos son *pattern search* ou algoritmos baseados no gradiente como o gradiente conxugado non lineal (NCG), BFGS, L- BFGS (descritos en [123], por exemplo) ou L- BFGS- B (ver [35], por exemplo). Doutra banda, cando consideramos problemas de optimización global, os algoritmos de optimización local anteriores poden terminar converxendo cara a mínimos locais, que non proporcionan a solución do problema plantexado. Este inconveniente pode ser resolto mediante a utilización de algoritmos de optimización estocásticos, por exemplo Simulated Annealing (SA). Aínda que a converxencia destes algoritmos de optimización estocásticos é lenta en xeral, teñen a vantaxe de que descartan os mínimos locais e evitan o quedar atrapados en solucións locais do problema.

Unha posibilidade para obter algoritmos de optimización global máis rápidos que foi considerada nesta tese é o uso de algoritmos híbridos, que combinan os algoritmos de optimización estocásticos cos métodos locais de optimización. Un exemplo desta clase de algoritmos é o algoritmo de Basin Hopping (BH), no cal o algoritmo global estocástico de SA emprégase para muestrear o espazo, xerando puntos veciños de maneira aleatoria, e combínase con algoritmos locais de tipo gradiente para capturar o mínimo partindo inicialmente dun deses veciños xerados pola mostraxe estocástica. Estes algoritmos analízanse nos traballos [78, 181]. Doutra banda, o custo computacional dos algoritmos resultantes xestionouse mediante unha paralelización eficiente nas arquitecturas de computadores axeitadas.

En relación con estes últimos aspectos computacionais, presentouse un estudo numérico para analizar o comportamento das metodoloxías propostas cando se aplican a unha carteira grande de contratos de seguros asociados a tipos de xuro, é dicir, os contratos pagan unha suma de beneficios en caso de falecemento do posuidor da póliza, sempre que isto ocorra antes dun tempo determinado. Nesta aplicación, úsase un modelo de mercado (LIBOR Market Model, LMM) para simular a evolución estocástica dos tipos forward no tempo. O LMM utilízase moito na práctica, pois presenta a vantaxe de poder ser calibrado a mercado de modo consistente.

O contido desta tese doutoral organizouse en dúas partes.

A Parte I contén as achegas que se levaron a cabo no marco da teoría moderna de medidas de risco e de estimación robusta do risco.

No Capítulo 2 revísanse brevemente os resultados de converxencia que motivan os problemas que se abordan nos seguintes capítulos. Ponse especial énfase na noción de topoloxía feble e nas formas comúns de simetría en probabilidade, por exemplo a estacionaridade (stationarity) e a intercambiabilidade (exchangeability).

No Capítulo 3 preséntase unha revisión particular da teoría moderna de medidas de risco, baseada fundamentalmente nas referencias que se citan neste capítulo. Dedicouse unha atención especial ao concepto de robustez cualitativa en estimación do risco.

O Capítulo 4 baséase fundamentalmente nos contidos do traballo [75], excepto a Sección 4.1, que non forma parte dese traballo e que representa o desenvolvemento de diferentes resultados da literatura, que se atoparon axeitados por razóns de completitude do capítulo. Neste capítulo introdúcese unha noción refinada de robustez cualitativa que se deriva a partir do concepto de consistencia estatística en termos de medidas de probabilidade aleatoria.

A Parte II da tese trata do problema de representación de carteiras, con atención especial ás que aparecen no sector de seguros de vida.

No Capítulo 5 recóllense os resultados fundamentais da teoría de integración estocástica en espazos de Banach, cunha certa revisión sobre a literatura no tema. Esta revisión baséase nas referencias citadas neste capítulo. Aínda que a maior parte dos resultados son coñecidos, na Sección 5.4 analizáronse e demostráronse aqueles que non se atoparon na forma axeitada na literatura.

Os Capítulos 6 e 7 inclúen resultados orixinais baseados no traballo [79]. En concreto, no Capítulo 6 abórdase o problema de representación de carteiras no marco da teoría de integración estocástica en espazos de Banach, mentres que no Capítulo 7 analízase esta formulación cando se tratan carteiras de seguros de vida, facendo especial énfase en carteiras compostas de pólizas de seguros de vida (whole life insurance policies).

O Capítulo 8 baséase no traballo [76]. En concreto, contén un estudo numérico da aproximación exposta no capítulo anterior cando se aplica a carteiras de seguros a prazo, considerando o LMM como modelo para describir a evolución no tempo da estrutura de tipos de xuro forward.

Finalmente, inclúese un apartado no que se analizan algúns problemas abertos e expone ideas de posibles liñas futuras de investigación.

Bibliography

[1] Carlo Acerbi. Spectral measures of risk: a coherent representation of subjective risk aversion. *Journal of Banking & Finance*, 26(7):1505–1518, 2002.

[2] Carlo Acerbi. Coherent representations of subjective risk aversion. *Risk Measures for the XXI Century*, 2003.

[3] Carlo Acerbi and Dirk Tasche. Expected shortfall: a natural coherent alternative to value at risk. *Economic notes*, 31(2):379–388, 2002.

[4] Carlo Acerbi and Dirk Tasche. On the coherence of Expected Shortfall. *Journal of Banking & Finance*, 26(7):1487–1503, 2002.

[5] David J Aldous. *Exchangeability and related topics*. Springer, 1985.

[6] Charalambos D Aliprantis and Kim Border. *Infinite dimensional analysis: a hitchhiker's guide*. Springer Science & Business Media, 2006.

[7] Philippe Artzner, Freddy Delbaen, Jean-Marc Eber, and David Heath. Coherent measures of risk. *Risk management: value at risk and beyond*, page 145, 2002.

[8] Philippe Artzner, Freddy Delbaen, and Pablo Koch-Medina. Risk measures and efficient use of capital. *Astin Bulletin*, 39(01):101–116, 2009.

[9] Pauline Barrieu and Giacomo Scandolo. Assessing financial model risk. *European Journal of Operational Research*, 242(2):546–556, 2015.

[10] Flavia Barsotti, Xavier Milhaud, and Yahia Salhi. Lapse risk in life insurance: Correlation and contagion effects among policyholders' behaviors. *Insurance: Mathematics and Economics*, 71:317–331, 2016.

[11] Daniel Bauer, Matthias Börger, and Jochen Ruß. On the pricing of longevity-linked securities. *Insurance: Mathematics and Economics*, 46(1):139–149, 2010.

[12] Fabio Bellini, Bernhard Klar, Alfred Müller, and Emanuela Rosazza Gianin. Generalized quantiles as risk measures. *Insurance: Mathematics and Economics*, 54:41–48, 2014.

[13] Denis Belomestny and Volker Krätschmer. Central limit theorems for law-invariant coherent risk measures. *Journal of Applied Probability*, 49(1):1–21, 2012.

[14] Patrizia Berti and Pietro Rigo. A Glivenko-Cantelli theorem for exchangeable random variables. *Statistics & probability letters*, 32(4):385–391, 1997.

[15] Enrico Biffis and Pietro Millossovich. A bidimensional approach to mortality risk. *Decisions in Economics and Finance*, 29(2):71–94, 2006.

[16] Patrick Billingsley. *Convergence of probability measures.* John Wiley & Sons, 2013.

[17] BIS. Minimum capital requirements for market risk, Bank for International Settlements. 2016.

[18] BIS. Revirion to Basel II market risk framework. Basel Committee on Banking Supervision, Bank for International Settlements. February 2011.

[19] David Blake, Tom Boardman, and Andrew Cairns. Sharing longevity risk: Why governments should issue longevity bonds. *North American Actuarial Journal*, 18(1):258–277, 2014.

[20] David Blake, Andrew Cairns, Guy Coughlan, Kevin Dowd, and Richard MacMinn. The new life market. *Journal of Risk and Insurance*, 80(3):501–558, 2013.

[21] Vladimir I. Bogachev. *Measure theory*, volume 2. Springer Science & Business Media, 2007.

[22] Vladimir I. Bogachev. *Measure theory*, volume 1. Springer Science & Business Media, 2007.

[23] Jean Bourgain. Some remarks on Banach spaces in which martingale difference sequences are unconditional. *Arkiv för Matematik*, 21(1-2):163, 1983.

[24] Thomas Breuer and Imre Csiszár. Measuring distribution model risk. *Mathematical Finance*, 26(2):395–411, 2016.

[25] Damiano Brigo and Fabio Mercurio. *Interest rate models-theory and practice: with smile, inflation and credit*. Springer Science & Business Media, 2007.

[26] Natacha Brouhns, Michel Denuit, and Jeroen K Vermunt. A Poisson log-bilinear regression approach to the construction of projected lifetables. *Insurance: Mathematics and economics*, 31(3):373–393, 2002.

[27] Z Brzeźniak and Annie Millet. On the stochastic Strichartz estimates and the stochastic nonlinear Schrödinger equation on a compact riemannian manifold. *Potential Analysis*, 41(2):269–315, 2014.

[28] Zdzisław Brzeźniak. Stochastic partial differential equations in M-type 2 Banach spaces. *Potential Analysis*, 4(1):1–45, 1995.

[29] Zdzisław Brzeźniak. On stochastic convolution in Banach spaces and applications. *Stochastics: An International Journal of Probability and Stochastic Processes*, 61(3-4):245–295, 1997.

[30] Zdzisław Brzeźniak. Some remarks on Itô and Stratonovich integration in 2-smooth Banach spaces. In *Probabilistic methods in fluids*, pages 48–69. World Scientific, 2003.

[31] Zdzislaw Brzeźniak, Jan M.A.M. van Neerven, Mark C. Veraar, and Lutz Weis. Itô's formula in UMD Banach spaces and regularity of solutions of the Zakai equation. *Journal of Differential Equations*, 245(1):30–58, 2008.

[32] Donald L. Burkholder. A geometrical characterization of Banach spaces in which martingale difference sequences are unconditional. *The Annals of Probability*, pages 997–1011, 1981.

[33] Donald L Burkholder. *Handbook of the geometry of Banach spaces*, volume 1. Elsevier, 2001.

[34] Donald L. Burkholder. Martingales and singular integrals in Banach spaces. *Handbook of the geometry of Banach spaces*, 1:233–269, 2001.

[35] Richard H. Byrd, Peihuang Lu, Jorge Nocedal, and Ciyou Zhu. A limited memory algorithm for bound constrained optimization. *SIAM J. on Scientific Computing*, 16:1190–1208, 1995.

[36] Andrew JG Cairns, David Blake, and Kevin Dowd. Pricing death: Frameworks for the valuation and securitization of mortality risk. *ASTIN Bulletin: The Journal of the IAA*, 36(1):79–120, 2006.

[37] Andrew JG Cairns, David Blake, and Kevin Dowd. A two-factor model for stochastic mortality with parameter uncertainty: theory and calibration. *Journal of Risk and Insurance*, 73(4):687–718, 2006.

[38] Andrew JG Cairns, David Blake, Kevin Dowd, Guy D Coughlan, David Epstein, Alen Ong, and Igor Balevich. A quantitative comparison of stochastic mortality models using data from England and Wales and the United States. *North American Actuarial Journal*, 13(1):1–35, 2009.

[39] Rene Carmona, Michael Tehranchi, et al. A characterization of hedging portfolios for interest rate contingent claims. *The Annals of Applied Probability*, 14(3):1267–1294, 2004.

[40] René Carmona and Michael R Tehranchi. *Interest rate models: an infinite dimensional stochastic analysis perspective*. Springer Science & Business Media, 2007.

[41] Patrick Cheridito and Tianhui Li. Risk measures on Orlicz hearts. *Mathematical Finance*, 19(2):189–214, 2009.

[42] Erhan Çınlar. *Probability and stochastics*, volume 261. Springer Science & Business Media, 2011.

[43] Rama Cont. Model uncertainty and its impact on the pricing of derivative instruments. *Mathematical finance*, 16(3):519–547, 2006.

[44] Rama Cont, Romain Deguest, and Giacomo Scandolo. Robustness and sensitivity analysis of risk measurement procedures. *Quantitative Finance*, 10(6):593–606, 2010.

[45] Antonio Cuevas. Qualitative robustness in abstract inference. *Journal of statistical planning and inference*, 18(3):277–289, 1988.

[46] Iain D Currie. Smoothing and forecasting mortality rates with p-splines. *Talk given at the Institute of Actuaries*, 2006.

[47] Iain D. Currie, Maria Durban, and Paul H.C. Eilers. Smoothing and forecasting mortality rates. *Statistical modelling*, 4(4):279–298, 2004.

[48] Mikkel Dahl. Stochastic mortality in life insurance: market reserves and mortality-linked insurance contracts. *Insurance: mathematics and economics*, 35(1):113–136, 2004.

[49] Rose-Anne Dana. A representation result for concave Schur functions. *Mathematical finance*, 15(4):613–634, 2005.

[50] Atul Dar and J. Colin Dodds. Interest rates, the emergency fund hypothesis and saving through endowment policies: some empirical evidence for the UK. *Journal of Risk and Insurance*, pages 415–433, 1989.

[51] Angelos Dassios and Hongbiao Zhao. A dynamic contagion process. *Advances in Applied Probability*, 43(3):814–846, 2011.

[52] Bruno de Finetti. Sull'approssimazione empirica di una legge di probabilità. *Giornale dell'Istituto Italiano degli Attuati*, 4(3):415–420, 1933.

[53] Bruno de Finetti. La prévision: ses lois logiques, ses sources subjectives. In *Annales de l'institut Henri Poincaré*, volume 7, pages 1–68, 1937.

[54] Bruno de Finetti. La probabilitá e la statistica nei rapporti con l'induzione, secondo i diversi punti di vista. In *Induzione e Statistica*, pages 1–122. Springer, 2011.

[55] Griselda Deelstra, Martino Grasselli, and Christopher Van Weverberg. The role of the dependence between mortality and interest rates when pricing guaranteed annuity options. *Insurance: Mathematics and Economics*, 71:205–219, 2016.

[56] Freddy Delbaen. Coherent risk measures on general probability spaces. In *Advances in finance and stochastics*, pages 1–37. Springer, 2002.

[57] Darinka Dentcheva, Spiridon Penev, and Andrzej Ruszczyński. Kusuoka representation of higher order dual risk measures. *Annals of Operations Research*, 181(1):325–335, 2010.

[58] Michel Denuit, Pierre Devolder, and Anne-Cécile Goderniaux. Securitization of longevity risk: Pricing survivor bonds with Wang transform in the Lee-Carter framework. *Journal of Risk and Insurance*, 74(1):87–113, 2007.

[59] Nils Detering and Natalie Packham. Model risk of contingent claims. *Quantitative Finance*, 16(9):1357–1374, 2016.

[60] Egbert Dettweiler. On the martingale problem for Banach space valued stochastic differential equations. *Journal of Theoretical Probability*, 2(2):159–191, 1989.

[61] Egbert Dettweiler. Stochastic integration relative to Brownian motion on a general Banach space. *Doga Mat*, 15(2):58–97, 1991.

[62] Richard M. Dudley. *Uniform central limit theorems*, volume 23. Cambridge Univ Press, 1999.

[63] Richard M. Dudley. *Real analysis and probability*, volume 74. Cambridge University Press, 2002.

[64] Richard M. Dudley. Distances of probability measures and random variables. *Selected Works of RM Dudley*, pages 28–37, 2010.

[65] Richard M. Dudley, Evarist Giné, and Joel Zinn. Uniform and universal Glivenko-Cantelli classes. *Journal of Theoretical Probability*, 4(3):485–510, 1991.

[66] EBA. Discussion paper on draft regultory technical standards on prudent valuation, under Article 100 of the draft Capital Requirements Regulation (CRR). Discussion Paper, European Banking Authority, November 2012.

[67] Gerald A Edgar and Louis Sucheston. *Stopping times and directed processes*, volume 47. Cambridge University Press, 1992.

[68] EIOPA. QIS5 Technical Specification. *European commission.*, 2010.

[69] Ivar Ekeland and Erik Taflin. A theory of bond portfolios. *The Annals of Applied Probability*, 15(2):1260–1305, 2005.

[70] Martin Eling and Michael Kochanski. Research on lapse in life insurance: what has been done and what needs to be done? *The Journal of Risk Finance*, 14(4):392–413, 2013.

[71] Paul Embrechts and Marius Hofert. Statistics and quantitative risk management for banking and insurance. *Annual Review of Statistics and Its Application*, 1:493–514, 2014.

[72] Susanne Emmer, Marie Kratz, and Dirk Tasche. What is the best risk measure in practice? A comparison of standard measures. *ESSEC Working Paper*, 2013.

[73] Ryszard Engelking. *General Topology*, volume 47. Polish Sci. Publ., 1992.

[74] Margarida de Faria, Maria João Oliveira, and Ludwig Streit. A generalized Clark-Ocone formula. *Random Operators and Stochastic Equations*, pages 163–174, 2000.

[75] José Luis Fernández, Enrico Ferri, and Carlos Vázquez. Asymptotic stability of empirical processes and related functionals. *Preprint. arXiv:1710.07070*, 2017.

[76] Ana María Ferreiro, Enrico Ferri, José Antonio García-Rodríguez, and Carlos Vázquez. Global optimization for model points automatic selection in life insurance portfolios. *Preprint*, 2018.

[77] Ana María Ferreiro, José Antonio García-Rodríguez García-Rodrǵuez, José Germán López-Salas, and Carlos Vázquez. An efficient implementation of parallel simulated annealing algorithm in GPUs. *Journal of Global Optimization*, 57(3):863–890, 2013.

[78] Ana María Ferreiro, José Antonio García-Rodríguez, Carlos Vázquez, Eliana Costa e Silva, and Aldina Correia. GPU parallelization of two-phase optimization algorithms. *Mathematics and Computers in Simulation*, 2018.

[79] Enrico Ferri. Optimal model points portfolio in life insurance. *Preprint. arXiv:1808.00866*, 2018.

[80] Damir Filipovic. *Consistency problems for Heath-Jarrow-Morton interest rate models*. Springer Science & Business Media, 2001.

[81] Tobias Fissler, Johanna F Ziegel, et al. Higher order elicitability and Osband's principle. *The Annals of Statistics*, 44(4):1680–1707, 2016.

[82] Hans Föllmer and Alexander Schied. Convex measures of risk and trading constraints. *Finance and stochastics*, 6(4):429–447, 2002.

[83] Hans Föllmer and Alexander Schied. *Stochastic finance: an introduction in discrete time*. Walter de Gruyter, 2011.

[84] Robert Fortet and Edith Mourier. Convergence de la répartition empirique vers la répartition théorique. In *Annales scientifiques de l'École Normale Supérieure*, volume 70, pages 267–285, 1953.

[85] David A. Freedman. Invariants under mixing which generalize de Finetti's theorem. *The Annals of Mathematical Statistics*, 33(3):916–923, 1962.

[86] Marco Frittelli and Emanuela Rosazza Gianin. Law invariant convex risk measures. In *Advances in Mathematical Economics*, pages 33–46. Springer, 2005.

[87] Peter Gänssler and Winfried Stute. Empirical processes: a survey of results for independent and identically distributed random variables. *The Annals of Probability*, pages 193–243, 1979.

[88] David John Haldane Garling. Brownian motion and UMD-spaces. In *Probability and Banach spaces*, pages 36–49. Springer, 1986.

[89] Izrail Moiseevich Gel'Fand and Z Jakubski. Generalized random processes. *Dokl. Akad. Nauk SSSR (N.S.)*, (100):853–856, 1955.

[90] Jayanta Kumar Ghosh and R.V. Ramamoorthi. *Bayesian Nonparametrics*. Springer, 2003.

[91] Paul Glasserman and Xingbo Xu. Robust risk measurement and model risk. *Quantitative Finance*, 14(1):29–58, 2014.

[92] Valery Glivenko and Francesco Paolo Cantelli. Sulla determinazione empirica della legge di probabilitá. *Giorn. Ist. Ital. Attuari. 4*, pages 92–99, 1933.

[93] Tilmann Gneiting. Making and evaluating point forecasts. *Journal of the American Statistical Association*, 106(494):746–762, 2011.

[94] Benjamin Gompertz. XXIV. On the nature of the function expressive of the law of human mortality, and on a new mode of determining the value of life contingencies. In a letter to Francis Baily, Esq. FRS &c. *Philosophical transactions of the Royal Society of London*, 115:513–583, 1825.

[95] Axel Grorud and Étienne Pardoux. Intégrales Hilbertiennes anticipantes par rapport à un processus de Wiener cylindrique et calcul stochastique associé. *Applied Mathematics and Optimization*, 25(1):31–49, 1992.

[96] Leonard Gross. Abstract Wiener spaces. *Proc Fifth Berkeley Sympos Math Statist and Probability*, 2:29–63, 1965.

[97] Frank R. Hampel. Contributions to the theory of robust estimation. *Unpublished dissertation, Univ. California, Berkeley.*

[98] Frank R. Hampel. A general qualitative definition of robustness. *The Annals of Mathematical Statistics*, pages 1887–1896, 1971.

[99] Erik Heitfield. Parameter uncertainty and the credit risk of collateralized debt obligations. *Board of Governors of the Federal Reserve System Working Paper*, 2009.

[100] Edwin Hewitt and Kenneth A Ross. Abstract harmonic analysis. *Bull. Amer. Math. Soc. 73 (1967), 292-294*, pages 0002–9904, 1967.

[101] Edwin Hewitt and Leonard J Savage. Symmetric measures on Cartesian products. *Transactions of the American Mathematical Society*, 80(2):470–501, 1955.

[102] Peter J Huber and Elvezio M Ronchetti. Robustness of design. *Robust Statistics, Second Edition*, pages 239–248, 1975.

[103] Luka Jalen and Rogemar Mamon. Valuation of contingent claims with mortality and interest rate risks. *Mathematical and Computer Modelling*, 49(9-10):1893–1904, 2009.

[104] Hans Jarchow. *Locally convex spaces*. Springer Science & Business Media, 2012.

[105] Philippe Jorion. Risk management lessons from the credit crisis. *European Financial Management*, 15(5):923–933, 2009.

[106] Olav Kallenberg. Spreading and predictable sampling in exchangeable sequences and processes. *The Annals of Probability*, pages 508–534, 1988.

[107] Olav Kallenberg. *Foundations of modern probability*. Springer Science & Business Media, 1997.

[108] Olav Kallenberg. *Probabilistic symmetries and invariance principles*. Springer Science & Business Media, 2006.

[109] Jeroen Kerkhof, Bertrand Melenberg, and Hans Schumacher. Model risk and capital reserves. *Journal of Banking & Finance*, 34(1):267–279, 2010.

[110] Paul C Kettler, Frank Proske, and Mark Rubtsov. Sensitivity with respect to the yield curve: duration in a stochastic setting. In *Inspired by Finance*, pages 363–385. Springer, 2014.

[111] John F.C. Kingman. Uses of exchangeability. *The Annals of Probability*, pages 183–197, 1978.

[112] Pablo Koch-Medina and Cosimo Munari. Law-invariant risk measures: Extension properties and qualitative robustness. *Statistics & Risk Modeling*, 31(3-4):215–236, 2014.

[113] Steven Kou, Xianhua Peng, and Chris C Heyde. External risk measures and Basel accords. *Mathematics of Operations Research*, 38(3):393–417, 2013.

[114] Volker Krätschmer, Alexander Schied, and Henryk Zähle. Qualitative and infinitesimal robustness of tail-dependent statistical functionals. *Journal of Multivariate Analysis*, 103(1):35–47, 2012.

[115] Volker Krätschmer, Alexander Schied, and Henryk Zähle. Comparative and qualitative robustness for law-invariant risk measures. *Finance and Stochastics*, 18(2):271–295, 2014.

[116] Volker Krätschmer, Alexander Schied, and Henryk Zähle. Quasi-Hadamard differentiability of general risk functionals and its application. *Statistics & Risk Modeling*, 32(1):25–47, 2015.

[117] Weiyu Kuo, Chenghsien Tsai, and Wei-Kuang Chen. An empirical study on the lapse rate: the cointegration approach. *Journal of Risk and Insurance*, 70(3):489–508, 2003.

[118] Daniel Lacker. Law invariant risk measures and information divergences. *arXiv preprint arXiv:1510.07030*, 2015.

[119] Daniel Lacker. Liquidity, risk measures, and concentration of measure. *Mathematics of Operations Research*, 43(3), 2018.

[120] Alexandra Lauer and Henryk Zähle. Nonparametric estimation of risk measures of collective risks. *Statistics & Risk Modeling*, 32(2):89–102, 2016.

[121] Jorge A Le, David Nualart, et al. Stochastic evolution equations with random generators. *The Annals of Probability*, 26(1):149–186, 1998.

[122] Ronald D. Lee and Lawrence R. Carter. Modeling and forecasting US mortality. *Journal of the American statistical association*, 87(419):659–671, 1992.

[123] Dong C. Liu and Jorge Nocedal. On the limited memory BFGS method for large scale optimization. *Math. Progr. B*, 45:503–528, 1989.

[124] Xiaoming Liu, Rogemar Mamon, and Huan Gao. A generalized pricing framework addressing correlated mortality and interest risks: a change of probability measure approach. *Stochastics An International Journal of Probability and Stochastic Processes*, 86(4):594–608, 2014.

[125] Stéphane Loisel, Pierre Arnal, and Romain Durand. Correlation crises in insurance and finance, and the need for dynamic risk maps in ORSA. *Working paper*, 2010.

[126] Stéphane Loisel and Xavier Milhaud. From deterministic to stochastic surrender risk models: Impact of correlation crises on economic capital. *European Journal of Operational Research*, 214(2):348–357, 2011.

[127] Jan Maas. Malliavin calculus and decoupling inequalities in Banach spaces. *Journal of Mathematical Analysis and Applications*, 363(2):383–398, 2010.

[128] Jan Maas and Jan M.A.M. van Neerven. A Clark-Ocone formula in UMD Banach spaces. *Electronic Communications in Probability*, 13:151–164, 2008.

[129] George W Mackey. Borel structure in groups and their duals. *Transactions of the American Mathematical Society*, 85(1):134–165, 1957.

[130] Paul Malliavin. Stochastic calculus of variations and hypoelliptic operators. *Stochastic Analysis, Kinokuniya, Tokyo*, pages 195–263, 1978.

[131] Eddy Mayer-Wolf and Moshe Zakai. The Clark–Ocone formula for vector valued Wiener functionals. *Journal of Functional Analysis*, 229(1):143–154, 2005.

[132] Eddy Mayer-Wolf and Moshe Zakai. The divergence of Banach space valued random variables on Wiener space. *Probability theory and related fields*, 132(2):291–320, 2005.

[133] Leslie Mayhew and David Smith. Human survival at older ages and the implications for longevity bond pricing. *North American Actuarial Journal*, 15(2):248–265, 2011.

[134] Terry R. McConnell. Decoupling and stochastic integration in UMD Banach spaces. *Probab. Math. Statist*, 10(2):283–295, 1989.

[135] Moshe A. Milevsky and S. David Promislow. Mortality derivatives and the option to annuitize. *Insurance: Mathematics and Economics*, 29(3):299–318, 2001.

[136] Ivan Mizera. Qualitative robustness and weak continuity: the extreme unction. *Nonparametrics and Robustness in Modern Statistical Inference and Time Series Analysis: A Festschrift in honor of Professor Jana Jurecková*, 1:169, 2010.

[137] Gaspard Monge. Mémoire sur la théorie des déblais et des remblais. *In Histoire de l'Académie des Sciences de Paris, avec les Mémoires de mathématique et de physique pour la même année*, pages 666–704, 1781.

[138] Massimo Morini. *Understanding and Managing Model Risk: A practical guide for quants, traders and validators.* John Wiley & Sons, 2011.

[139] Arnold L. Neidhardt. *Stochastic integrals in 2-uniformly smooth Banach spaces.* PhD thesis, Ph. D. Thesis, University of Wisconsin, 1978.

[140] David Nualart. *The Malliavin calculus and related topics*, volume 1995. Springer, 2006.

[141] Martin Ondreját. *Equations d'évolution stochastiques dans les espaces de Banach: unicités abstraites, propriété de Markov forte, équations hyperboliques.* PhD thesis, Nancy 1, 2003.

[142] Kent Osband. *Providing incentives for better cost forecasting.* PhD thesis, University of California, Berkeley, 1985.

[143] Horst Osswald. On the Clark Ocone formula for the abstract Wiener space. *Advances in Mathematics*, 176(1):38–52, 2003.

[144] Kalyanapuram Rangachari Parthasarathy. *Probability measures on metric spaces*, volume 352. American Mathematical Soc., 1967.

[145] Gilles Pisier. Martingales with values in uniformly convex spaces. *Israel Journal of Mathematics*, 20(3-4):326–350, 1975.

[146] Marcin Pitera and Thorsten Schmidt. Unbiased estimation of risk. *arXiv preprint arXiv:1603.02615*, 2016.

[147] Chris Preston. Some notes on standard Borel and related spaces. *arXiv preprint arXiv:0809.3066*, 2008.

[148] Yu V. Prokhorov. Convergence of random processes and limit theorems in probability theory. *Theory of Probability & Its Applications*, 1(2):157–214, 1956.

[149] Matthijs Pronk and Mark Veraar. Tools for Malliavin calculus in UMD Banach spaces. *Potential Analysis*, 40(4):307–344, 2014.

[150] Svetlozar T. Rachev. The Monge-Kantorovich mass transference problem and its stochastic applications. *Theory of Probability & Its Applications*, 29(4):647–676, 1985.

[151] Svetlozar T. Rachev and Ludger Rüschendorf. *Mass Transportation Problems: Volume I: Theory*, volume 1. Springer Science & Business Media, 1998.

[152] R. Ranga Rao. Relations between weak and uniform convergence of measures with applications. *The Annals of Mathematical Statistics*, pages 659–680, 1962.

[153] Arthur E. Renshaw and Steven Haberman. Lee–Carter mortality forecasting with age-specific enhancement. *Insurance: Mathematics and Economics*, 33(2):255–272, 2003.

[154] Federal Reserve. Supervisiory guidance on model risk management. Board of Governors of the Federal Reserve System, Office of the Comptroller of the Currency, SR Letter 11-7 Attachement, April 2011.

[155] Markus Riedle. Cylindrical Wiener processes. In *Séminaire de Probabilités XLIII*, pages 191–214. Springer, 2011.

[156] Nathanael Ringer and Michael Tehranchi. Optimal portfolio choice in the bond market. *Finance and Stochastics*, 10(4):553–573, 2006.

[157] R. Tyrrell Rockafellar. *Conjugate duality and optimization*, volume 16. SIAM, 1974.

[158] Andrzej Ruszczyński and Alexander Shapiro. Optimization of convex risk functions. *Mathematics of operations research*, 31(3):433–452, 2006.

[159] Czesław Ryll-Nardzewski. On stationary sequences of random variables and the de Finetti's equivalence. In *Colloquium Mathematicae*, volume 4, pages 149–156. Institute of Mathematics Polish Academy of Sciences, 1957.

[160] Hiroshi Satô. Gaussian measure on a Banach space and Abstract Winer measure. *Nagoya Mathematical Journal*, 36:65–81, 1969.

[161] Irving E. Segal. Tensor algebras over Hilbert spaces. I. *Transactions of the American Mathematical Society*, 81(1):106–134, 1956.

[162] Alexander Shapiro. On Kusuoka representation of law invariant risk measures. *Mathematics of Operations Research*, 38(1):142–152, 2013.

[163] Alexander Shapiro and Darinka Dentcheva. *Lectures on stochastic programming: modeling and theory*, volume 16. SIAM, 2014.

[164] Yang Shen and Tak Kuen Siu. Longevity bond pricing under stochastic interest rate and mortality with regime-switching. *Insurance: Mathematics and Economics*, 52(1):114–123, 2013.

[165] S. Sinai, Ya. *Dynamical System II: Ergodic Theory with Applications to Dynamical Systems and Statistical Mechanics.* Springer-Verlag Berlin Heidelberg, 1989.

[166] Yongsheng Song and Jia-An Yan. Risk measures with comonotonic subadditivity or convexity and respecting stochastic orders. *Insurance: Mathematics and Economics*, 45(3):459–465, 2009.

[167] Volker Strassen. The existence of probability measures with given marginals. *The Annals of Mathematical Statistics*, pages 423–439, 1965.

[168] Dirk Tasche. Expected shortfall and beyond. *Journal of Banking & Finance*, 26(7):1519–1533, 2002.

[169] Ramon van Handel. The universal Glivenko–Cantelli property. *Probability Theory and Related Fields*, pages 1–24, 2013.

[170] Jan van Neerven, Mark C. Veraar, and Lutz Weis. Stochastic maximal Lp-regularity. *The Annals of Probability*, 40(2):788–812, 2012.

[171] Jan M.A.M. van Neerven. Stochastic evolution equations. *ISEM lecture notes*, 2008.

[172] Jan M.A.M. van Neerven. γ-Radonifying operators–a survey. *arXiv preprint arXiv:0911.3788*, 2009.

[173] Jan M.A.M. van Neerven, Mark C. Veraar, and Lutz Weis. Conditions for stochastic integrability in UMD Banach spaces. *Banach spaces and their applications in analysis (in honor of Nigel Kalton's 60th birthday)*, pages 127–146, 2007.

[174] Jan M.A.M. van Neerven, Mark C. Veraar, and Lutz Weis. Stochastic integration in UMD Banach spaces. *The Annals of Probability*, 35(4):1438–1478, 2007.

[175] Jan M.A.M. van Neerven, Mark C. Veraar, and Lutz Weis. Stochastic evolution equations in UMD Banach spaces. *arXiv preprint arXiv:0804.0932*, 2008.

[176] Jan M.A.M. van Neerven, Mark C. Veraar, and Lutz Weis. Stochastic maximal Lp-regularity. *The Annals of Probability*, 40(2):788–812, 2012.

[177] Jan M.A.M. van Neerven and Lutz Weis. Stochastic integration of functions with values in a Banach space. *Studia Mathematica*, 166:131–170, 2005.

[178] Jan M.A.M. van Neerven and Lutz Weis. Weak limits and integrals of Gaussian covariances in Banach spaces. *Probability and mathematical statistics*, 25(1):55, 2005.

[179] Veeravalli S. Varadarajan. On the convergence of sample probability distributions. *Sankhyā: The Indian Journal of Statistics (1933-1960)*, 19(1/2):23–26, 1958.

[180] Cedric Villani. Topics in optimal transportation (graduate studies in mathematics, vol. 58). *American Mathematical Soc.*, (58), 2003.

[181] David J. Wales and Jonathan P. K. Doye. Global optimization by Basin-Hopping and the lowest energy structures of Lennard-Jones clusters containing up to 110 Atoms. *J. Phys. Chem. A*, 101:5111–5116, 1997.

[182] Stefan Weber. Distribution-invariant risk measures, information, and dynamic consistency. *Mathematical Finance*, 16(2):419–441, 2006.

[183] Henryk Zähle. Marcinkiewicz–Zygmund and ordinary strong laws for empirical distribution functions and plug-in estimators. *Statistics*, 48(5):951–964, 2014.

[184] Henryk Zähle. Qualitative robustness of von Mises statistics based on strongly mixing data. *Statistical Papers*, 55(1):157–167, 2014.

[185] Henryk Zähle. Qualitative robustness of statistical functionals under strong mixing. *Bernoulli*, 21(3):1412–1434, 2015.

[186] Henryk Zähle. A definition of qualitative robustness for general point estimators, and examples. *Journal of Multivariate Analysis*, 143:12–31, 2016.

[187] Yixing Zhao, Rogemar Mamon, and Huan Gao. A two-decrement model for the valuation and risk measurement of a guaranteed annuity option. *Econometrics and Statistics*, 8:231–249, 2018.

[188] Nan Zhu and Daniel Bauer. Applications of forward mortality factor models in life insurance practice. *The Geneva Papers on Risk and Insurance-Issues and Practice*, 36(4):567–594, 2011.

[189] Johanna F Ziegel. Coherence and elicitability. *Mathematical Finance*, 26(4):901–918, 2016.